AFIT/GSM/LSY/92S-5

AD-A258 989

GRAPHIC REPRESENTATION IN MANAGERIAL
DECISION MAKING: THE EFFECT OF SCALE BREAK
ON THE DEPENDENT AXIS

THESIS

Clark R. Carvalho, Captain, USAF

Michael D. McMillan, Captain, USAF

AFIT/GSM/LSY/92S-5

DTIC
ELECTE
DEC 17 1992
S B D

92-31542

1108

Approved for public release; distribution unlimited

92 12 16 022

The views expressed in this thesis are those of the authors and do *not reflect the official policy or position* of the Department of Defense or the U.S. Government.

DTIC QUALITY INSPECTED 2

Accession For	
NTIS GRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

AFIT/GSM/LSY/92S-5

GRAPHIC REPRESENTATION IN MANAGERIAL DECISION MAKING: THE EFFECT OF SCALE BREAK ON THE DEPENDENT AXIS

THESIS

Presented to the Faculty of the School of Systems and Logistics

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Systems Management

Clark R. Carvalho, B.S., M.S.

Captain, USAF

Michael D. McMillan, B.A.

Captain, USAF

September 1992

Approved for public release; distribution unlimited

Preface

The purpose of this study was to determine if a scale break on the dependent (vertical) axis of a graph had an effect on interpretation of the data presented. Decision makers are increasingly using data presented in a graphical format as a basis for their decisions. This data must be presented in a clear understandable format to facilitate timely accurate decision making.

An experiment was conducted, using a pen and paper format, to determine if decision makers derive different meanings from graphs with a scale break on the vertical axis. The experiment followed the pretest-posttest control group design. The control group was exposed to graphs constructed following the requirements for high integrity graphs. The experimental group was exposed to graphs that also followed high integrity graph criteria, except for a scale break on the dependent axis. By measuring the subject's response to the graphs, it was determined that a scale break on the dependent axis affects the interpretation of data presented in a graphical format.

We are deeply indebted to our thesis advisors, Major David Christensen and Mr. Richard Antolini, for their guidance, help, and support. We also wish to thank Dr. Guy Shane for the insight he provided on experimental methods. Finally, we would like to thank our families for their help, support, understanding, and love.

Clark R. Carvalho

Michael D. McMillan

Table of Contents

List of Figures

List of Tables

Page

AFIT/GSM/LSY/92S-5

Abstract

This thesis investigated whether a scale break on the dependent axis of a graph affects a decision maker's interpretation of the data presented in the graph. Tufte's lie factor was used to determine the level of distortion present in the graphs. A literature search revealed criteria for constructing high integrity graphs and formatting scale breaks. An experiment was conducted on 147 subjects to determine the effect of a scale break on the dependent axis. Graphs following the criteria for high integrity graphs were presented to the control group, while graphs following the criteria for high integrity graphs, with the exception of the scale break, were presented to the experimental group. Using a parametric two-sample t test and a non-parametric Rank Sum test, it was shown that data presented in a graph with a scale break is interpreted differently from data presented in a graph without a scale break. Analysis of variance (ANOVA) was conducted on the demographic factors for each subject. In the experimental group, sex and professional experience were factors that led to different interpretations of graphs with a scale break. The level of experience using graphs in decision making was also a factor that led to different interpretations of the graphs in both groups.

GRAPHIC REPRESENTATION IN MANAGERIAL DECISION MAKING:
THE EFFECT OF SCALE BREAK ON THE DEPENDENT AXIS

I. Introduction

General Issue

The graphical display of numeric data is becoming more widespread and a common feature of modern decision support systems. This phenomenon is a result of the increased use of micro-computer based, graphically oriented software packages and the perceived benefits of a graphical versus tabular format for data presentation. Previous research, however, has shown that poorly constructed graphs can result in a misinterpretation of the trends or indications present in the underlying data (Larkin, 1990; Kern, 1991; Taylor, 1983). For this reason, a common set of standards for graph construction is required. In addition, the standards should be empirically based, using the accuracy of information communicated as a criterion.

Specific Problem

The use of graphically displayed information has become a standard practice in modern business decision situations. This trend is based on the results of studies which indicate the potential benefits of visual presentations, as well as the ease of use of current graphical software packages. Studies have shown that combining graphics with verbal presentations helps to get a point across better and faster, helps a group reach consensus, and helps with the retention of ideas (Howard, 1988:93). At the same time, software packages have been introduced which allow anyone with the ability to operate a personal computer to create high quality business graphics (Howard, 1988: 92). So, it is easy to see why decision makers rely on graphically displayed data.

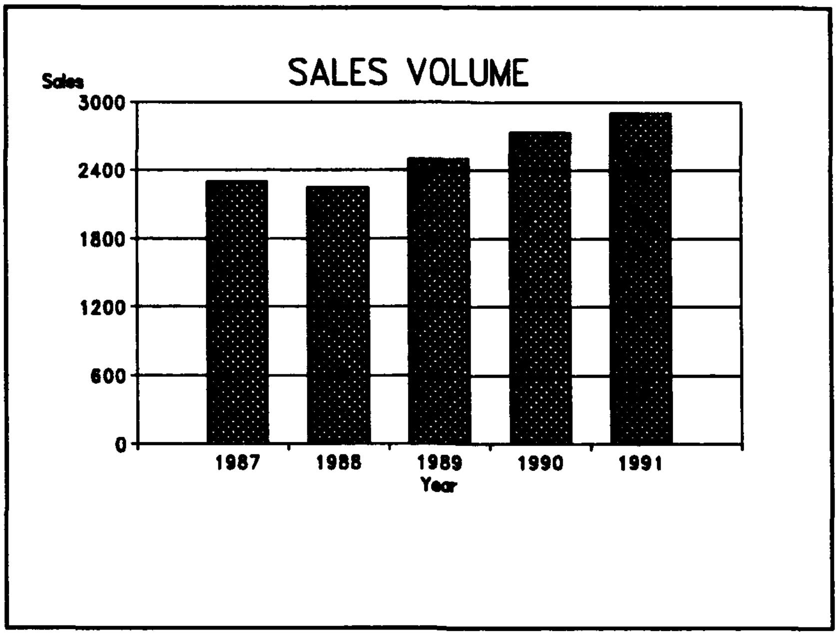
The average decision maker has a wide variety of software applications with graphical capabilities at his disposal, both spreadsheets applications and business presentation applications. All of these programs, to some degree, simplify the process of construction a graph from complex data. They also provide the user with a tremendous amount of flexibility in determining the final appearance of any graph produced. A recent review of business graphics applications indicated that 90% allowed manual scaling of the axis and over 70 percent possessed drawing tools (Howard, 1988:176; Seymour, 1988:95). However, previous research has shown that manipulations which alter the appearance of a graph can also have an impact on the message a decision maker derives from the graph (Larkin, 1990; Kern, 1991; Taylor, 1983). To prevent the inadvertent (or intentional) misrepresentation of data through a graphical presentation, both the person constructing the graph and the person interpreting it, must reference a common set of standards for high integrity graphs.

Standards which address the construction of graphs have been in existence for quite some time. For example, the *Journal of the American Statistical Association* published guidelines as early as 1915 (Joint Committee on Standards for Graphic Representation, 1915). But, as the popularity of graphical data presentation has increased, so has the number and variety of standards for their construction increased. Christensen and Larkin point out that "some of these guidelines are more concerned with style", while others are concerned with the integrity of the graph (Christensen, 1992: 131). The issue of a graph's integrity directly affects whether a decision maker is likely to be misled by data presented in the graph (Larkin, 1990: 58).

While previous researchers investigated specific criteria which impact the integrity of a graph, the vast majority of the existing standards contain criteria which have yet to receive empirical scrutiny. Of particular interest are the criteria related to the construction of scaling for the dependent (typically the vertical) axis. Most standards are in agreement with the notion that the dependent axis scale should include a zero point. No doubt, this high level of agreement is a result of an almost universal appreciation of the significance effect produced by not including a full range of values in a graph. If a full range of values is not included, the graph does not accurately represent the ratios present in the underlying data. Tufte addresses this point and suggests that "the representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numeric quantities represented" (Tufte, 1983: 56). In addition, he proposes a "lie factor" which is designed to quantify the degree of any misrepresentation caused by arbitrary scaling of the dependent axis.

There are two techniques for omitting selected portions of the dependent variable range. One method is to truncate the scale at the lower end of values, which is another way of saying "omit the zero point" as mentioned above. This method was investigated by Kern and found to be misleading (Kern, 1991:38-39). A second method to omit data values would be to include the zero point, but delete some intermediate portion of the remaining values. This second method is advocated by several authors through the use of a scale break (Rogers, 1969; Schmid and Schmid, 1954).

Any interruption of the dependent axis scaling will induce a misrepresentation according to Tufte's lie factor. Figure 1 gives an example of the visual distortion created when the dependent axis in not continuous from a zero base line. However, there may be situations where the full representation of data values reduces the resolution of information of the data to the point of being meaningless (Cleveland, 1988:79). It is possible that a properly formatted scale break will adequately signal to the decision maker the absence of full representation in the graph and emphasize the relevant range of the data. If the decision maker is aware of this inconsistency in the graph, he may be able to adequately compensate and correctly interpret the information conveyed by the graph. The issue of continuous (no



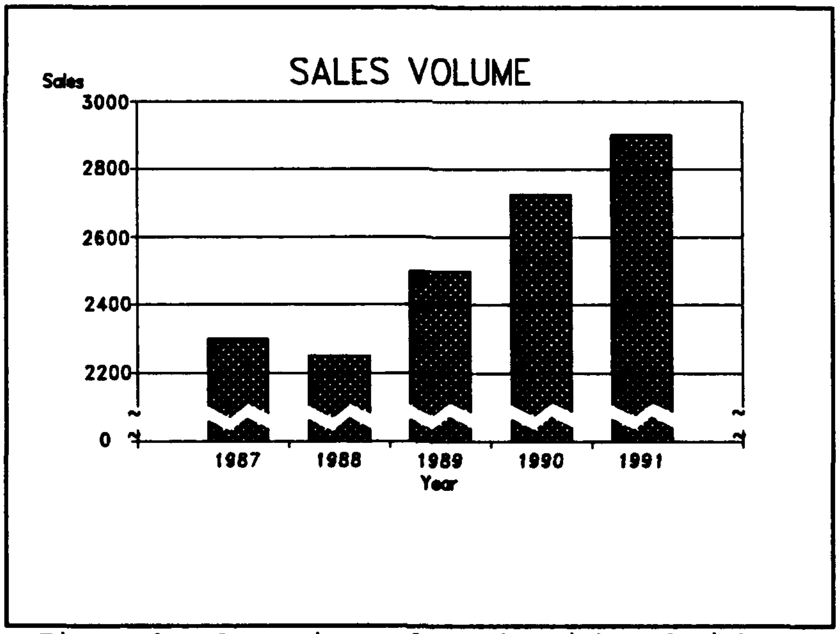


Figure 1. Comparison of Graphs with and without Scale Break

scale break) versus non-continuous (scale break present) scaling of the dependent axis is the focus of this research.

The specific hypothesis to be tested is:

Null Hypothesis: There is no difference between the interpretation of graphs with and without a break in the scale of the dependent axis.

Investigative Questions

To adequately investigate this hypothesis, the following investigative questions will be addressed:

1. What are the existing standards involving scale breaks? Are these standards empirically grounded?

2. How can a scale break be drawn or constructed using popular software applications?

3. What are the managerial implications of graphs containing scale breaks?

4. Are graphs with a scale break on the dependent (vertical) axis interpreted differently from graphs without a scale break?

5. Does the magnitude of the distortion, as measured by Tufte's lie factor, produced by a scale break affect the interpretation of the graph?

6. Are line graphs with a scale break more likely to be misinterpreted than bar graphs with a scale break?

7. Are there any demographic factors which affect the interpretation of graphs with scale breaks?

Limitations

This thesis contains several limitations which reduce the scope of the research. The focus of the research effort was to evaluate the possible differences in interpretation of graphs with and without scale breaks on the dependent axis. The decision making tasks involved in this evaluation were very limited. Subjects were asked to respond with either agreement or disagreement to a statement describing information presented in a graph. Their responses were based on their impression generated from viewing the graphs for a short period of time (approximately 15 seconds). So, the decision task was a simple matter of indicating a level of agreement with a proposed conclusion. A second limitation involves the types of graphs evaluated. Only two types of graphs, vertical bar and line graphs, were evaluated. The decision to limit the research to these two graph types was based on the appropriateness of these two types in decision support situations and their susceptibility to manipulation by breaking the dependent axis scale. Additional limitations, specific to the experimental design, will be discussed in Chapter III, Methodology.

Conclusion

Investigative questions 1 through 3, as well as a discussion of other literature relevant to this research, is presented in Chapter II, Literature Review. The remainder of the investigative questions are addresses through a carefully designed and executed behavioral experiment. Chapter III, Methodology, covers the specifics of the experimental design, its execution, and the statistical manipulations required for analysis. The results of the experiment are discussed in Chapter IV, Analysis and Findings. The final chapter, Chapter V, Conclusion, contains a summary and proposed interpretation of the experimental findings, as well as recommendations for further research.

II. Literature Review

This literature review has a twofold purpose, to provide answers to the first three investigative questions and to provide a sound basis for understanding the motivation and methodology of this research. It will examine the general climate as it relates to business graphics and the implications of their use. It will also address the body of knowledge which attempts to define how graphs should be properly constructed.

The chapter is divided into five sections. The first section covers the uses of graphics in business decision making and the impact misinterpretation of graphic data can have on the decision making process. The next area of discussion is previous research and findings in the area of misleading graphs. The following two sections define the currently proposed "standards" for the construction of graphs and, more specifically, scale breaks. The final section describes the capabilities of current graphics software packages to display scale breaks and the specific procedures used to draw a scale break using a spreadsheet application (Quattro Pro).

Current Uses of Business Graphics

The current trend in business decision making situations is to turn more and more frequently to graphs as a means of conveying information and making a point. In fact, for many decision makers the credibility and business judgment of those who does not use high quality, professional graphics as part of their presentations comes into question (Seymour, 1988:94). The availability of high quality, flexible business graphics software and the personal computers on which to run it have been the major drivers in this development. The flexibility available in current decision support systems allows decision makers to "turn their financial spread sheets into colorful graphs or extract rich graphical representations of information in existing databases"

(Jarvenpaa, 1989:285). However, a decision support system's flexibility also presents an obstacle.

The vast number of graphical format options available to decision makers makes it difficult for them to decide which format is the most effective for their situation. The effectiveness of a particular graphic representation is dependent on the characteristics of the decision problem at hand (Jarvenpaa, 1989:285). So, it becomes a critical issue that the designers of business graphics recognize the characteristics of the decision task and the graph format which is most effective in that situation.

The effect of scale breaks on the graph formatting decision can be inferred from Cleveland's concern for resolution in graphical representation (Cleveland, 1985:79). As Cleveland points out, there often are situations where a significant change or trend present in the data is lost in a graph displaying the full range of data values. Such a situation would exit in a system which operates at a high level of a variable, in absolute terms, but a small percentage change in the level of the variable has a dramatic effect. An example of such a system would be a firm which has a high volume of sales but a small profit margin. A decision maker in this environment would need to detect very small changes in sales volume before profits are adversely affected. In this situation a scale break would allow the graphics designer to limit the range of values displayed so that small changes in sales volume are accentuated. But, accuracy is not the only aspect of decision making affected by formatting.

Much of the past research into the issue of decision making based on graphic representation has focused on task accuracy, and the results have been mixed. Jarvenpaa, however, points out that there are additional implications of graphic decision making. He concluded that since the format of a graph is closely tied to the decision task, "changes in a presentation format can lead to changes in the decision strategy used" (Jarvenpaa, 1989:298). A manifistation of this correlation of graph format and decision style was longer decision times when improperly formatted graphs were used.

While "graphical information presentation may be the most effective means for facilitating comparative analysis, pattern finding, and sequencing activities (Carey, 1991:78)," the medium is not without its pitfalls. The sheer frequency of graph use combined with the flexibility of formatting options makes the potential for faulty graphical decision making a real possibility. As Tan points out:

> The use of graphical packages have reached a stage where it is now both practical and profitable to train designers and end-users on how to identify situations in which a particular display alternative may be more or less appropriate. (Tan, 1990:417)

Previous Research into the Subject of Misleading Graphs

As critical as the format of a graph may be to the accurate communication of information, only a limited amount of work has been accomplished to evaluate formatting criteria. A review of the literature uncovered three research efforts into the subject of misleading graphs. These efforts show a progression from a general indication that graphs can be misleading to an evaluation of the contribution made by specific principles of graph construction. Each of the studies focused on the decision tasks associated with a specific population of test subjects. While this would tend to limit the generalizability of the resultant findings, the overall impression left by these works is that the manipulation of a graph's form can change the message it conveys. Some of the specifics of each of the reviewed research efforts are discussed in the following paragraphs.

One of the earliest works in this area is the research conducted by Taylor in 1983. Taylor focused on the impact financial graph format manipulations had on the message perceived by the users of the graphs. She was further interested in identifying which of the evaluated manipulations produced the greatest response. The study involved an evaluation of the responses of a group of bank loan officers to graphs portraying the financial position of selected firms. For experimental purposes the graphs were manipulated in violation of the following "caveats" of graph construction:

> 1. Scale range - Don't extend the range very much beyond the highest or lowest points unless you are sure the results will be a more realistic picture.
>
> 2. Grid proportions - Contracting or expanding either or both the vertical and horizontal scales can radically alter the configuration of the curves and consequently convey entirely different visual impressions.
>
> 3. Zero-based point of reference - The omission of zero magnifies changes and may make unimportant changes seem important.
>
> 4. Semilogarithmic scale - Rate of change charts should not be used for public presentations.
>
> 5. Strata charts - Generally, the stratum exhibiting marked irregularities should be placed at or near the top of the graph.
>
> 6. Multiple-amount scales - Multiple-amount scales should be used with caution or misrepresentation of relationships is likely to occur.
>
> 7. Presentation of declining profits - The discretionary selection of years to be presented may affect a viewer's perceptions.
>
> 8. Financial statement order - The financial statement order of presenting time for the horizontal scale of the graph may create a different illusion of company performance. (Taylor, 1983:13-21)

After viewing the graphs, the loan officers were then asked to rate the financial risk of the firms strictly on the basis of the graphical presentations. An analysis of the experimental results led Taylor to state that "The potential for manipulating user's perceptions of financial graphs is great unless both preparers and viewers of graphs are aware of potentially misleading formats" (Taylor, 1983:117). In particular, of the eight caveats evaluated, five were found to produce significant levels of misinterpretation: zero-based point of reference, semilogarithmic scaling, multiple-amount scaling, discretionary use of years presented, and presentation of data in financial statement order (Taylor, 1983:117).

A study conducted by Larkin in 1990 had an intent similar to that of the Taylor study, but the structure was different. Larkin was interested in determining if graphs of Cost Performance Reports used in United States Air Force acquisition programs could be misleading if they were constructed in violation of specific criteria (Larkin, 1990:8-9). In addition to evaluating the responses of a different test population, the study considered different graph construction criteria from those evaluated by Taylor. The criteria tested by Larkin included:

> 1. The general arrangement of a graph should be from left to right and from bottom to top.
>
> 2. In strata (area) charts, the stratum with the least variability should be on the bottom.
>
> 3. Incorrect labels can create different impressions on users.
>
> 4. The number of dimensions in the graph should not exceed the number of dimensions in the data. (Larkin, 1990:36-37)

These criteria fall into a category which contribute to what Larkin terms "high integrity graphs:" graphs which faithfully present the information contained in the underlying data. The primary finding of this study was that low integrity graphs, those constructed in violation of one of the high integrity criteria, could mislead Air Force decision makers (Larkin, 1990:58). Additional findings were: (1) the graph types most frequently used in Cost Performance Reports are line and bar charts, and (2) most graphs were constructed with computer software, and (3) many of the graphs contained violations of high integrity criteria. A major contribution of the Larkin study, in addition to the experimental findings, was a synopsis of existing graph construction criteria.

A research effort by Kern also addressed whether graph construction could lead to misinterpretation, but the focus was much narrower. Kern focused on the effect Tufte's lie factor has on graphical representation. Tufte's lie factor is an attempt to quantify the amount of distortion present in a visual presentation, when compared to the underlying numeric data. Basically, the lie factor is a ratio of the amount of change present in the visual data compared to the amount of change present in the numeric data. The ratio can be written in the following form:

$$\text{Lie Factor} = \frac{\text{Size of Effect Shown in Graphic}}{\text{Size of Effect in Data}} \qquad (1)$$

Specifically, Kern was concerned with two questions. The first question asked if charts with a lie factor of greater than 1.05 or less than .95 could mislead decision makers. The second question was an attempt to correlate the level of misleading influence possessed by a graph with the magnitude of the graph's lie factor (Kern, 1991:6). To produce the desired lie factor in his experimental graphs, Kern manipulated the scale of the dependent axis so that it started at a point other than zero, a violation of a criterion evaluated by Taylor. Kern's findings supported those of Taylor; both positive and negative trend graphs with a lie factor outside the range of .95 to 1.05 were shown to be misleading. However, Kern was unable to establish any correlation between the level of the lie factor and the degree to which a graph was misinterpreted (Kern, 1991:38-39).

Criteria for Graph Construction

The literature provides a rich source of information regarding "standards" for the construction of graphs and charts. One of the earliest works, published in 1915 by the *Journal of the American Statistical Association*, contained simple guidelines which were suitable for their time period. However, as the use of graphs became more widespread, and correspondingly, the potential for their misuse became greater, the number and variety of published standards became more diverse. The guidance provided in these standards falls into two broad categories. The first of these categories has to do with the criteria necessary to construct a graph of high integrity. As was mentioned earlier, a high integrity graph is one which represents the underlying tabular data with a high degree of fidelity; the value of Tufte's lie factor for this type of graph would be very close to "1". The majority of the criteria in this group have to do with the scaling of the axis.

The second group of criteria can collectively be termed "style guides". They are concerned with techniques for graph construction which can have a significant effect on the clarity or effectiveness of a graph in conveying the information contained in the underlying data. A violation of one of these criteria may not have as dramatic an effect on the interpretation of the graph as one of the high integrity criteria.

A weakness inherent in the majority of the published criteria is the lack of empirical justification. Of the 15 sources reviewed, only three provided empirical backing for their position (Kern, 1991; Larkin, 1990; Taylor, 1983). In addition, for many criteria there is disagreement between the various authors as to the importance or advisability of a particular criterion. These factors make any attempt to synthesize all of the sources into a single format difficult. However, Larkin produced a valuable synopsis of the existing criteria in his 1990 research by reducing the combined list of standards into a matrix referenced by author (Larkin, 1990:21-24). This format is recreated in Appendix A with minor modifications.

One of the more interesting features of the existing graphical formatting criteria is disagreement on including zero in the scale. While the majority of authors recommend the inclusion of zero, Cleveland states that "the need for zero is not so compelling that we should allow its inclusion to ruin the resolution of the data in the graph" (Cleveland, 1985:76). His position is based on the assumption that a critical reader will analyze the scale tick mark labels and understand their implications.

A similar disagreement exists over the issue of scale break usage. The authorities are fairly equally split on whether or not to condone formatting a graph so that there is a break in the dependent axis. The Joint Committee on Standards for Grapic Representation, and MacGregor both favor the use of scale breaks when a large portion of the dependent axis grid is unnecessary (Joint Committee on Standards for Graphic Representation, 1915:92; MacGregor, 1979:23-24). Auger and Cleveland argue against the use of scale breaks which give inaccurate impressions (Auger, 1979:142; Cleveland, 1985:85). Two additional sources, the

American Society of Mechanical Engineers and Schmid and Schmid, give weight to both sides of the issue; these sources feel that the best format depends on the factors involved in the situation at hand (American Society of Mechanical Engineers, 1979:17-18; Schmid and Schmid, 1954:35). Unfortunately, none of the sources has the benefit of a empirical study as justification.

Formatting of Scale Breaks

The use of scale breaks presents a potential solution to both Cleveland's concern for resolution and the desire to include a zero-base. By starting a graph's vertical axis at zero and then breaking the scale to omit those portions of the scale range which contribute little to the graph's message, both positions can be satisfied. The question then becomes one of how to format the scale break so that a "critical reader" is aware of its existence.

Four of the authors reviewed offer suggested procedures for formatting scale breaks. These procedures are listed in Tables 1 and 2 in a tabular format with author references. Interestingly, Cleveland, who advocates omitting zero, offers only a "full scale break" as a means of depicting non-continuous scale. This, in effect, reduces the original chart to a series of separate charts with adjacent scaling (Cleveland, 1985:86). The results of this approach can be seen in Figure 2 for a bar graph and Figure 5 for a line graph. The remainder of the authors suggest procedures which would allow the data to be represented in a single graph. The most frequently recommended method for breaking the scale of a bar chart is to break the scale on the left and right of the chart, and show the break across each of the bars. For line charts the most common procedure is to include the zero-base and break the scale on both the left and right sides of the chart with a ragged or wavy line. Each of the suggested methods of formatting scale breaks is illustrated in Figures 2 through 8. All of the sources additionally suggest that a scale break on a line chart should not interrupt the plotted data.

TABLE 1

METHODS OF DRAWING SCALE BREAKS IN BAR CHARTS

AUTHOR (YEAR):	Use full scale break, no data connected across break.	Break excessively long bars beyond the next longest bar.	Break should be indicated across both left and right scale as well as all bars.
Cleveland (1985)	X		
Schmid (1983)		X	
Schmid and Schmid (1954)			X
Rogers (1961)			X

TABLE 2

METHODS OF DRAWING SCALE BREAKS IN LINE CHARTS

AUTHORS (YEAR):	Use a full scale break, no data connected across break.	Include the zero-base and show a ragged break across the entire chart.	Omit the zero-base and show the lower limit of the chart as a ragged line.	Include the zero-base and break the left and right vertical scales with a ragged line.
Cleveland (1985)	X			
Schmid (1983)				X
Schmid and Schmid (1954)		X	X	X
Rogers (1961)				X

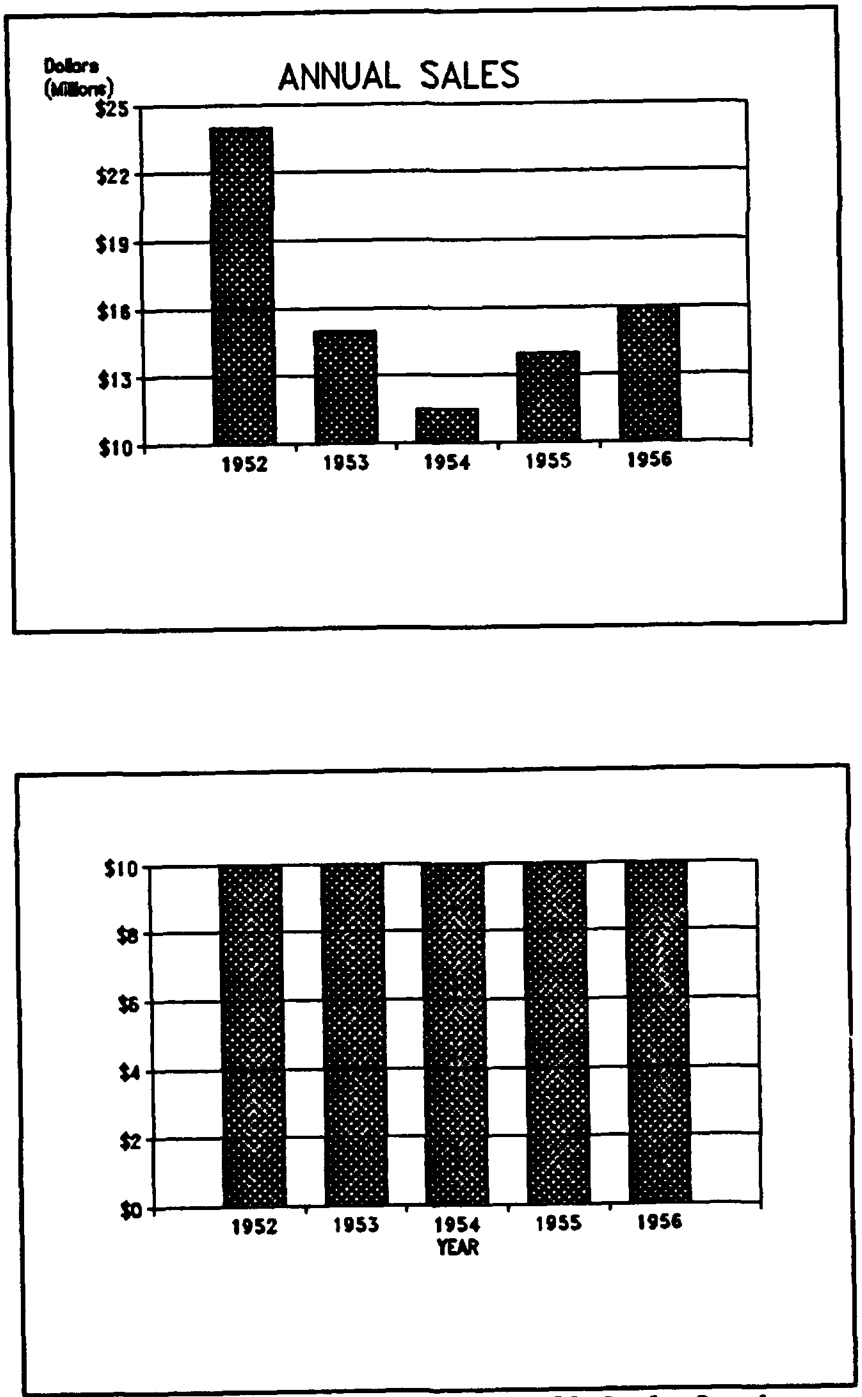


Figure 2. Bar Graph with Full Scale Break

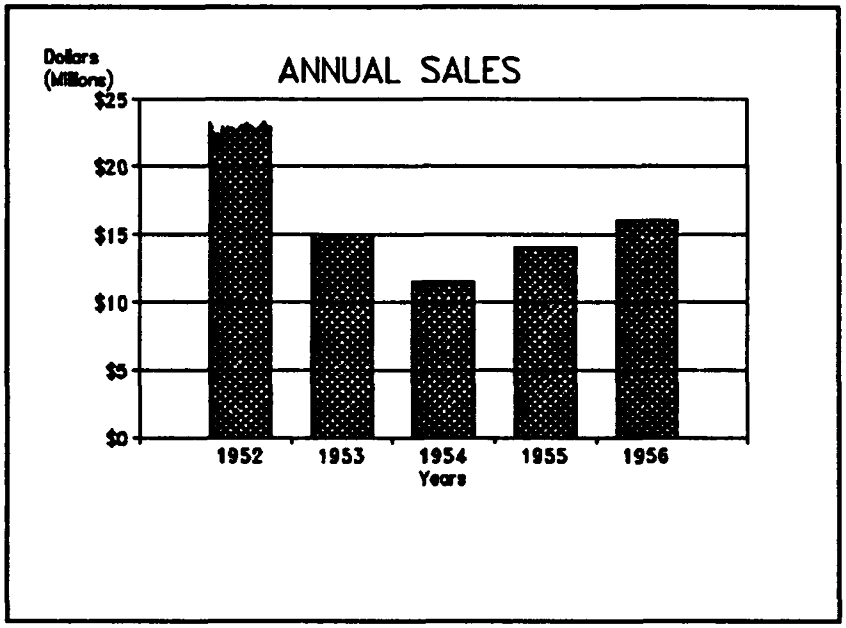


Figure 3. Bar Graph with Disproportionately Long Bar Broken

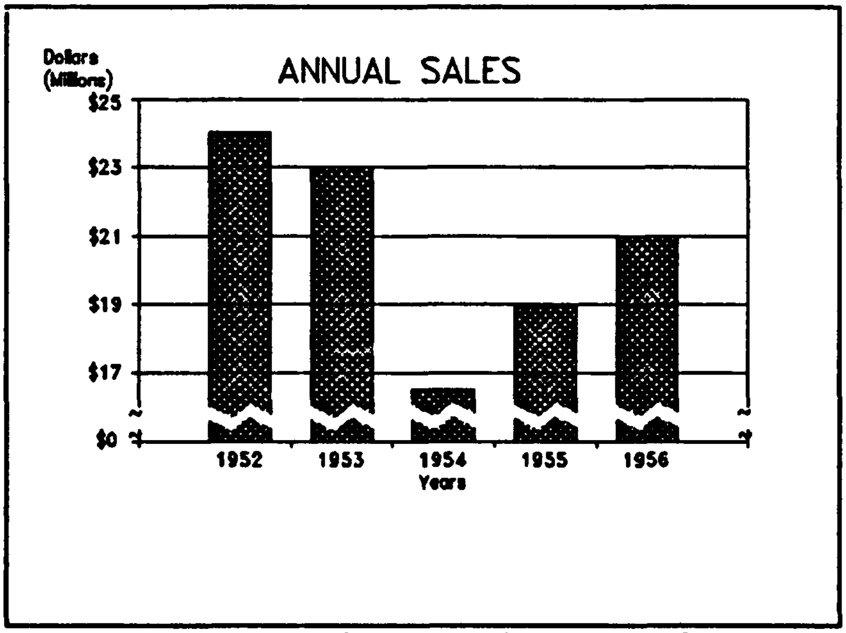


Figure 4. Bar Graph with Broken Vertical Scales and Broken Bars

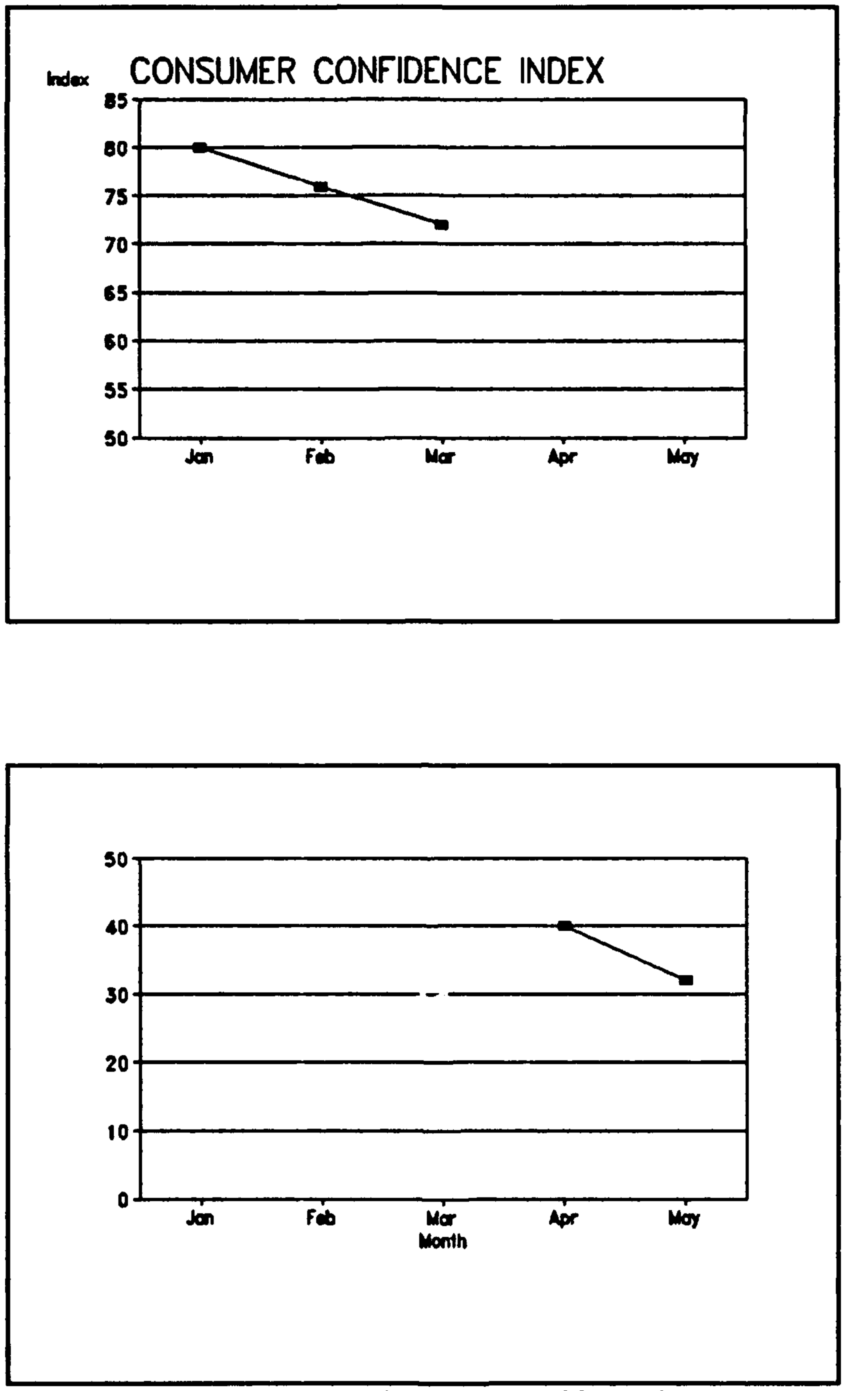


Figure 5. Line Graph with a Full Scale Break

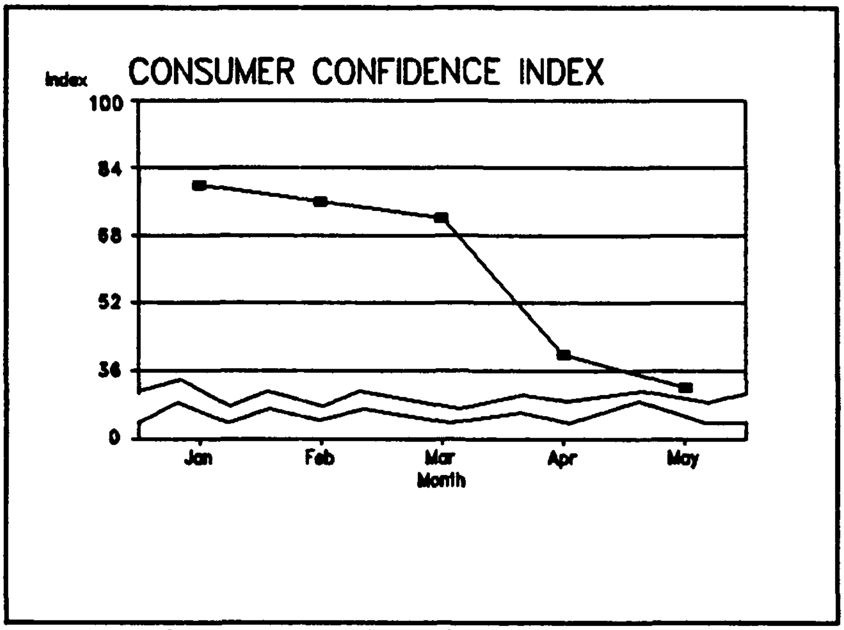


Figure 6. Line Graph with Zero Included, Scale Broken with Ragged Region Across Entire Graph

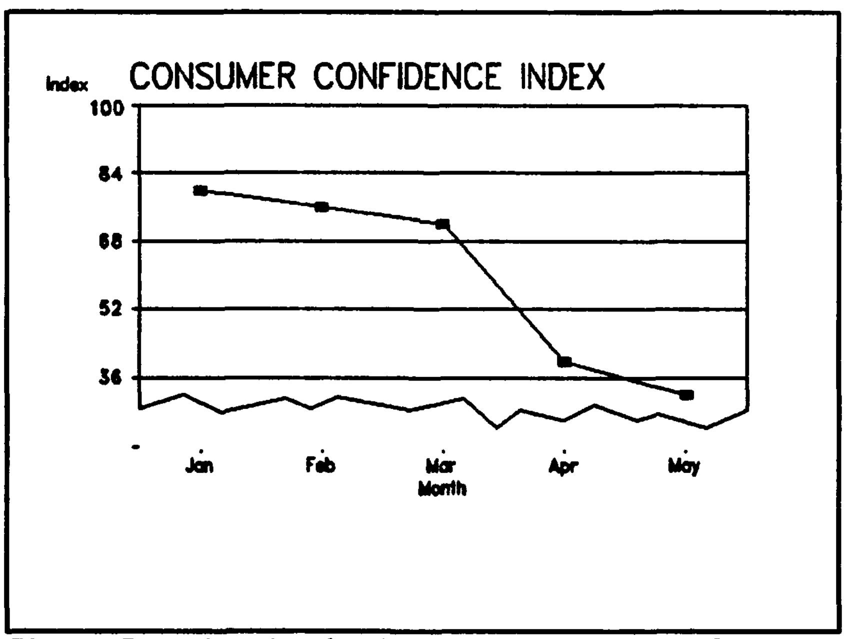


Figure 7. Line Graph with Zero Omitted, Scale Break Depicted by Ragged Bottom Edge

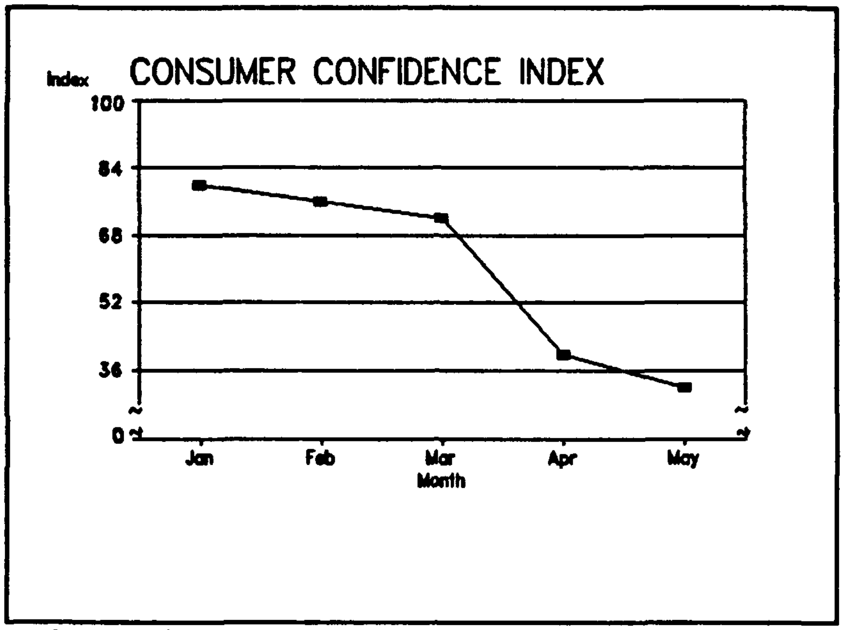


Figure 8. Line Graph with Zero Included, Scale Broken on Both Vertical Scales

Scale Break Creation Capabilities of Graphics Software

A review of the capabilities of current graphic software applications indicates that these packages have limited ability to create scale breaks in graphs. This conclusion is infered from the fact that none of the three comprehensive articles reviewed even addresses the issue of scale break (Howard, 1988; Seymour, 1988; GaCote, 1992). They do, however, give an indication of the level of flexibility possessed by the various graphics packages. Niney percent of the evaluated programs have the capability to manually adjust the scaling of the axis and over seventy percent contain drawing tools. Quattro Pro, a spreadsheet program, is representative of capabilities possessed by this family of software. For this reason, as well as ease of access, this program was adopted for the construction of graphs in this research. The procedures used by the researchers to construct scale breaks will now be discussed.

There is no standard feature in Quattro Pro specifically designed to break the scale of a graph. However, it is possible to construct a reasonable facsimile of a scale break for both line and bar charts with the "graph annotator" function. There are three basic steps required to design a graph with a scale break using this method: (1) specify the basic features of the graph, (2) modify the dependent axis scale, and (3) paste a suitably drawn "break" over the affected portions of the graph. For graphics programs other than Quattro Pro, unless the program being used is designed to create a scale break, this process will require trial and error to perfect.

The specific procedure used by the researchers to create the scale breaks for this experiment will now be described. The first step was to construct the graph using standard Quattro Pro features; this consisted of selecting the graph type and specifying the data. Once the basic graph has been constructed, the scaling on the dependent (y-axis) must be modified. This was done by selecting the "Y-Axis" feature under the "Graph" menu. The scale increment must be set to manual, and the high and low values as well as the desired increment must be specified. The "Display Scaling" feature must be turned off to prevent the system from displaying the full range of dependent variable values. The next step is to go to the "Annotate" feature of the graph menu. Because the Y-axis scale increment will not be displayed, the scale values must be defined for each tick mark. This is done by typing the appropriate value in a text box adjacent to the tick mark.

Although, the vertical axis scale can not actually be broken in Quattro Pro, it must appear this way on the graph. This is done by placing rectangular boxes, that are the same color and shading as the graph background, over the dependent axis on both the left and right sides. This gives the appearance of a gap in the axis lines. Once this is done, a tilde (~) is placed on top and bottom of the rectangular gap at each of the four points were the scale appears to be broken.

Since the bars in bar graphs extend upward from the zero line through the scale break, this graph type requires breaks for the bars themselves, as well as the scale. This break in the bar in effect splits the bar into two different sections. The procedures for accomplishing this step in Quattro Pro is very similar to the procedures used to break the scale. Using the "shape" function of the annotator, a box with jagged lines on the top and bottom is created. Once again, the shading of this jagged box must be the same as the graph background. Experience has shown that it is easier to draw these boxes in a large scale and then reduce the size to fit over the bar. This need only be accomplished once, because the remaining boxes can be copied and pasted where needed.

While the above procedures may appear cumbersome, they only need to be accomplished once for each graph type. By highlighting (selecting) the images that make up the specific graph scale break they may be saved in a clipboard file. This will allow the user to standardize the scale break once it is created, and then copy it onto other graphs when needed.

Summary

This literature review has answered the first three investigative questions as well as providing some background on the issues relevant to this research effort.

The importance of graphic representation in current business decision making situations was shown. Graphs are being used more frequently, and, as a result, the number of software applications which meet this need are becoming more numerous and capable. With the increased capability comes greater flexibility and the dilemma of determining the correct format for the graph.

Incorrect formatting was also shown to be a detrimental factor in graphical decision making; decision speed and accuracy can suffer when graphs are poorly drawn. An investigation of previous research revealed the potential misleading effects of improper graph formatting with respect to selected formatting criteria. The issue of breaking the scale of the dependent axis, however, has not been investigated.

The variety of sources which provided recommended "standards" for the construction of graphs was combined into a single comprehensive table. An analysis of this table showed general disagreement over two important issues, the inclusion of a zero base for the dependent axis and whether or not to include a scale break. Neither side in these disagreements has the weight of scientific study on their behalf.

Of the authors who condone the use of scale breaks, several provided techniques for actually formatting the break on the graph. A summary of these techniques was produced and the most frequent method identified for both line and bar charts. With these techniques in mind, a review of current graphics software revealed that the majority of the applications in use have limited capacity to produce scale breaks. However, the procedures which can be used to produce a scale break with a representative spreadsheet program, Quattro Pro were presented.

III. Methodology

This thesis is an extension of prior work on misleading graphics (Larkin, 1990; Kern, 1991). The primary objective is to determine if a scale break on the dependent (vertical) axis affects a decision maker's interpretation of data presented in a graphical format. The investigative questions are as follows:

1. What are the existing standards involving scale breaks? Are these standards empirically grounded?

2. How can a scale break be drawn or constructed using popular software applications?

3. What are the managerial implications of graphs containing scale breaks?

4. Are graphs with a scale break on the dependent (vertical) axis interpreted differently from graphs without a scale break?

5. Does the magnitude of the distortion, as measured by Tufte's lie factor, produced by a scale break affect the interpretation of the graph?

6. Are line graphs with a scale break more likely to be misinterpreted than bar graphs with a scale break?

7. Are there any demographic factors which affect the interpretation of graphs with scale breaks?

Investigative questions 1 through 3 were answered in Chapter II, Literature Review. A behavioral experiment using paper copies of computer generated graphics was undertaken to determine the answers to investigative questions 4 through 7. The specifics about how the experiment was designed, conducted, and analyzed will be covered in the following sections.

Experimental Design

Every experiment seeks to produce valid results. According to Emory there are two types of validity, internal and external. His explanation of these two types of validity is: "internal validity--do the conclusions we draw about a demonstrated experimental relationship truly imply cause?", and "external validity--does an observed causal relationship generalize across persons, settings and times? (Emory, 1991:424)." The seven major internal validity problems identified by Emory are:

1. History: While an experiment is taking place, some events may occur that confuse the relationship being studied.

2. Maturation: Changes that take place within the subject over time that are not specific to any particular event.

3. Testing: The process of taking a test affecting the scores of later tests.

4. Instrumentation: Changes between observations, in measuring instrument or observer.

5. Selection: The differential selection of subjects to be included in experimental and control groups.

6. Statistical Regression: The selection of study groups based on their extreme scores.

7. Experiment Mortality: Composition of the study groups change during the test. (Emory, 1991:424-426)

Emory states there are three threats to external validity:

1. The Reactivity of Testing on the Experimental Factor: Sensitizing subjects by the pretest so that they respond to the experimental stimulus in a different way.

2. Interaction of Selection and the Experimental Factor: The selection of test subjects; the population from which one actually selects may not be the same population one wishes to generalize to.

3. Other Reactive Factors: The experimental settings may bias a subject's response. (Emory, 1991:427)

This experiment used the pretest-posttest control group design. This design was selected because it does a good job of addressing the seven major internal validity problems encountered in experimentation (Emory, 1991:431). The threat of history was minimized as a result of the timing between the pretest and the posttest; the posttest was administered immediately after the pretest. Maturation was controlled by limiting the amount of time the subjects have to view each graph and by limiting the number of graphs. The effects of testing were expected to be minimal because the subjects only took the test once. Instrumentation was controlled by following a specified routine during each test. The random assignment of test subjects to either the control or experimental group should control the effects of selection, statistical regression, and experiment mortality.

While the pretest-posttest control group design strengthens internal validity, it does not do as good a job of controlling external validity. There is a chance for a reactive effect from testing (Emory, 1991:431). The pretest and posttest graphs had common characteristics to control for this reactive effect. These common characteristics were such things as the same number of bar and line graphs in each test, and designating half the graphs in each test for an "agree" conclusion and half for a "disagree" conclusion. Mask graphs were also included to reduce the reactivity effect. Additionally, all subjects were given the same initial graph (a mask) to anchor their responses.

The pretest-posttest control group design consists of two groups, control and experimental. In this case the control group was administered graphs that met all the requirements of high integrity graphs, while the experimental group was administered graphs that met all the requirements of high integrity graphs in the pretest, and graphs that violated one requirement for high integrity graphs, the broken vertical scale, in the posttest. *Subjects were randomly* (R) *assigned to* each of the groups. The diagram for this design is:

	PRETEST	MANIPULATION	POSTTEST	
R	O_1	X	O_2	(Experimental Group)
R	O_3		O_4	(Control Group)

The "R" in each group indicates a random selection of test subjects. The "X" is a treatment or manipulation of the independent variable. The "O" is an observation or measurement of the dependent variable (Emory, 1991:428-431). The effect of the experimental variable (E) is measured by the following relationship:

$$E = (O_2 - O_1) - (O_4 - O_3)$$

To answer investigative question 6, the following null (Ho) and alternative (Ha) hypotheses were developed:

$$Ho: \quad (O_2 - O_1) - (O_4 - O_3) = 0$$

$$Ha: \quad (O_2 - O_1) - (O_4 - O_3) \neq 0$$

The null hypothesis (Ho) states that a scale break on the dependent axis does not have an effect on interpretation of data represented in a graph. The alternative hypothesis (Ha) states that a scale break on the dependent axis does have an effect on graph interpretation. To address this hypothesis the difference between the pretest and posttest responses for each group must first be determined. Then the difference between the groups must be determined. It is expected that there will be little, if any, difference between the pretest responses for the two groups. The answers to investigative questions 4 through 7 were determined by conducting an analysis of variance (ANOVA) on the factors of interest. The statistical procedures used to analyze the hypothesis and conduct the ANOVA will be described in a later section.

Construction of the Experiment

There were three types of graphs used in the experiment: pretest, posttest, and mask. There were six graphs of each type, for a total of eighteen graphs. Each graph provided the subject with information in a graphical format and a conclusion with which they had to agree or disagree. The graphs were constructed using Quattro Pro, a spreadsheet program published by Borland. A nine-point Likert scale was used to gage the subjects' agreement or disagreement to the conclusions provided for each graph. According to the U.S. Army Research Institute for the Behavioral and Social Sciences the number of scale points used depends on the research design, the area of application, and the types of anchors used (Army Research Institute, 1989:119). A nine-point scale was selected in an effort to remain consistent with previous research into related subject matter (Taylor, 1983:42; Kern, 1991:26), and because it provided more response flexibility than a five or seven point scale.

Figure 9 contains two of the posttest graphs. The control graph (no scale break) is presented first, followed by the experimental graph (with scale break). Also shown in Figure 9 is the graph's conclusion statement and the Likert scale described above. Appendix B contains a complete set of the graphs used in the experiment.

The graphs used in the pretest followed the guidelines established for high integrity graphs. All of these graphs were developed using standard Quattro Pro graph settings; there was no attempt to distort the visual appearance of the graphs in any way. The lie factors associated with these graphs ranged from .949 to 1.04. The same pretest graphs were given to both the experimental and the control group. Three of the pretest graphs were bar graphs, and three were line graphs. Half of the graphs were designed with the conclusion statement worded so that the subject's response should agree (agree conclusion) with the conclusion statement. The other half of the graphs were designed with the conclusion statement worded so that the subject's response should disagree (disagree conclusion) with the conclusion statement. The design of half the graphs having an agree conclusion and the other half having a disagree conclusion was done to "safeguard against response-set bias" (Emory, 1991:221). Different descriptors (significantly, relatively, and about) were used to vary the wording of the conclusion statements in the graphs. This was done in an attempt to keep the subjects from getting bored. Table 3 is a summary of the features of each of the pretest graphs.

The graphs used in the posttest for the control group matched the pretest graphs by following the guidelines established for high integrity graphs. All of the control group graphs were designed using standard settings, and their lie factors ranged from .903 to 1.13. The graphs used in the posttest for the experimental group were identical to those for the control group (constructed from the same tabular data), except that all graphs for the experimental group had broken scales on the dependent axis (see Figure 9). The scale breaks were formatted in

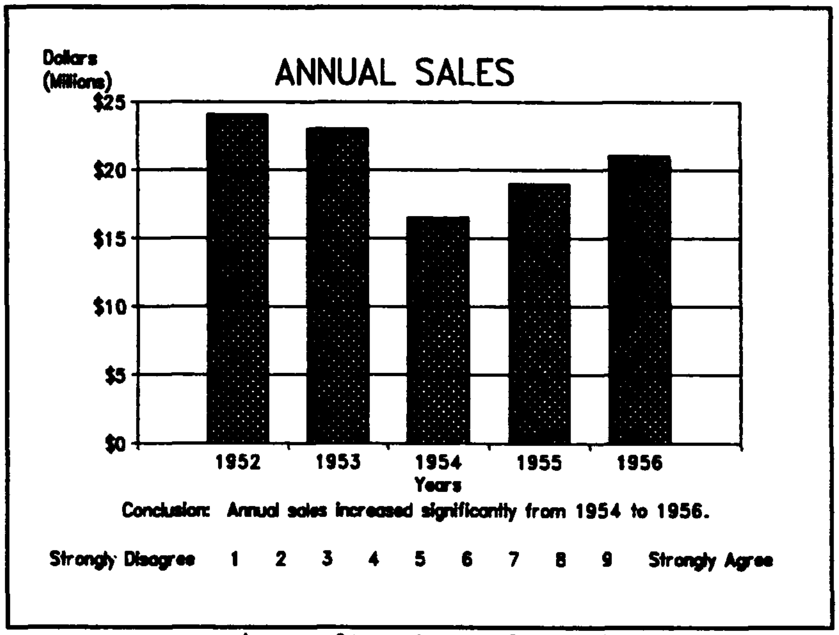


Figure 9A. Control Graph

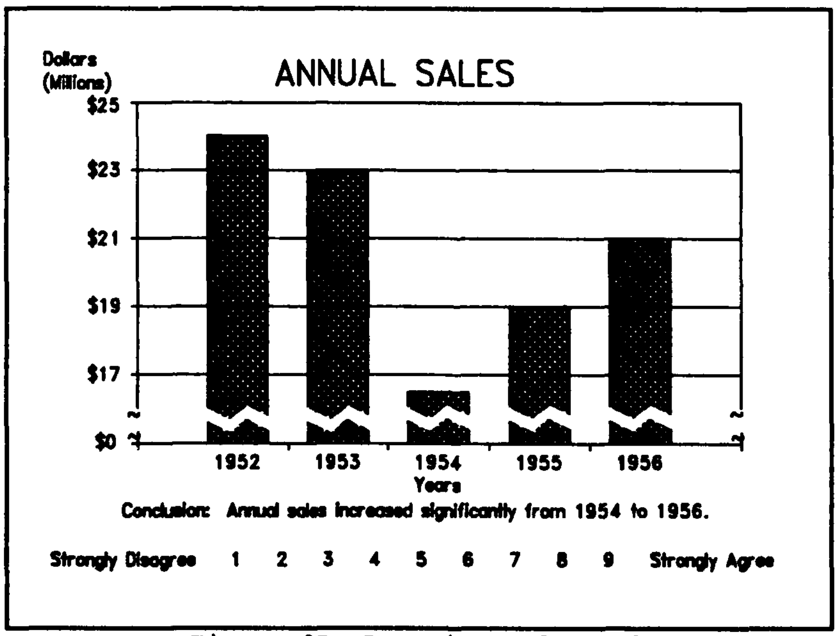


Figure 9B. Experimental Graph

compliance with the standards identified in Chapter II. On the bar graphs, the scale break splits both the vertical scaling lines and the bar into two distinct sections. On the line graphs, the scale break did not break the plotted line, only the vertical scaling lines on each side of the graph. Figure 10 shows a line graph with a scale break. The posttest graphs all had a mixture of features similar to those described above for the pretest graphs.

One additional feature, the lie factor, was manipulated in the experimental posttest graphs. Three of the experimental graphs were designed to have a dramatic break in the scale, and three graphs were designed to have nondramatic breaks in the scale. Graphs E1, E4, and E5 are categorized as dramatic, and graphs E2, E3, and E6 are categorized as nondramatic. This was done to provide two distinct levels of visual distortion in the experimental posttest graphs. The initial criterion used to determine the portion of dependent variable values omitted from the graph was the subjective evaluation of the researchers. The level of distortion was later quantified with a measure of Tufte's lie factor. The physical size of all of the scale breaks is the same within each graph type; it is the range of values represented by the dependent variable scale that changes from graph to graph. The features of the posttest graphs and the lie factors for each experimental graph are shown in Table 4.

The graphs used as masks included graphs distinguished from the pretest and posttest graphs in terms of graph type and task uniqueness. The same mask graphs were given to both the experimental and the control group. Masking graphs were interspersed throughout the pretest and posttest graphs for both the control and experimental groups. The purpose of the mask graphs was to reduce the reactivity of testing. By disguising the true purpose of the experiment from the subjects, the distinction between the pretest and posttest was blurred. This was accomplished by making the masks significantly different from the

TABLE 3

SUMMARY OF PRETEST GRAPH FEATURES

GRAPH #	DESCRIPTOR	CONCLUSION	GRAPH TYPE
1	remained relatively constant	disagree	bar
2	increased about	agree	bar
3	declined significantly	agree	line
4	increased significantly	disagree	line
5	remained relatively constant	agree	line
6	remained relatively constant	disagree	bar

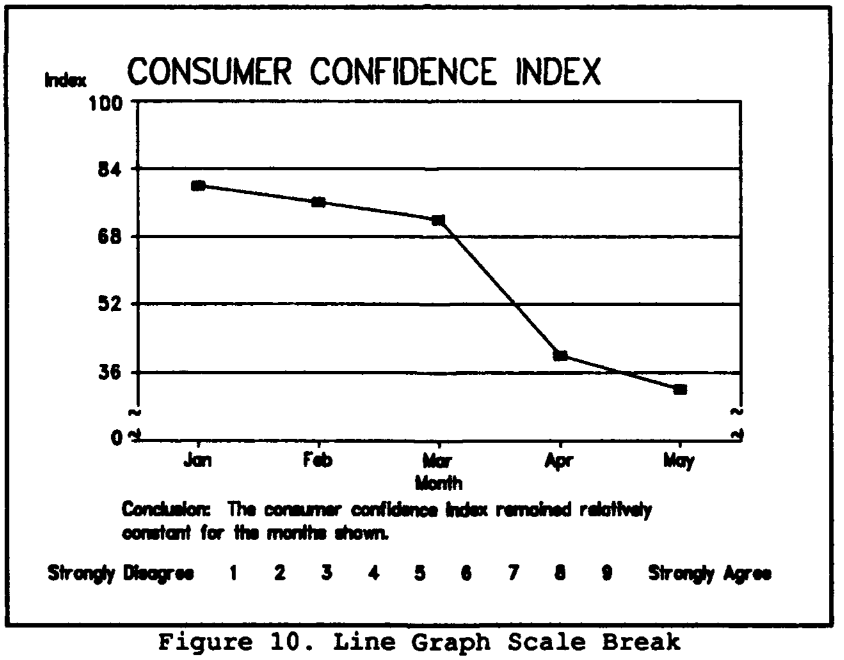


Figure 10. Line Graph Scale Break

pretest and posttest graphs. The features of the mask graphs are summarized in Table 5.

Conducting the Experiment

The experiment was conducted using a paper and pen format. This format was selected over a computer operated format because it reduced the reactive effects associated with experimental setting (experiment conducted in classrooms familiar to subjects) and media type. The paper and pen format also provided research flexibility and ease of administration. This format was adopted over a computer format due to the fact that the computer format would introduce other variables, such as computer literacy and computer speed and availability, that were not easily controlled.

Paper copies of the graphs were made, and the subjects indicated their response directly on the paper with a pencil or pen. Each test package consisted of two pages of instructions, six pretest graphs, six posttest graphs, six mask graphs, and a demographic questionnaire. Each package had the instructions on top followed by the same mask graph, VCR and TV Sales, which was used as an anchor. The anchor was used so that any uncertainty about the experiment and the effects of testing would be largely expended on a graph that was not scored (Shane:1992). Next came the first section of the test that consisted of the six pretest graphs and two of the mask graphs. This was followed by the second section consisting of the six posttest graphs and the three remaining mask graphs. The demographic questionnaire was the last item in the package.

The primary purpose of the demographic questionnaire was to identify personal factors which impacted the subjects' responses. The authors felt that familiarity and experience with graphic decision making would have the greatest effect on the interpretation of the graphs. As a result, each of the items included in the questionnaire, with the exception of age and sex, were selected based on a perceived

TABLE 4

SUMMARY OF POSTTEST GRAPH FEATURES

GRAPH #	DESCRIPTOR	CONCLUSION	GRAPH TYPE	EXP GRAPH LIE FACTOR
1	increased significantly	disagree	bar	12
2	significantly higher	agree	bar	6
3	remained relatively constant	disagree	line	3
4	fairly constant	agree	line	37
5	beginning to stabilize	agree	line	9
6	fairly constant	agree	bar	3

TABLE 5

SUMMARY OF MASK GRAPH FEATURES

GRAPH #	DESCRIPTOR	CONCLUSION	GRAPH TYPE
1	remained relatively constant	agree	2-line
2	falls significantly	disagree	line
3	increased faster	agree	2-bar
4	generated the majority	agree	pie
5	outside the specified tolerance	disagree	p-chart
6	remained consistent	disagree	2-line

relationship to previous use of graphs in decision making. The demographic factors the authors selected are:

1. Sex
2. Age
3. Education Level
4. Primary Field of Professional Experience
5. Years of Federal Employment
6. How Often Subject Uses Graphs Decision Making
7. How Often Subject Constructs Graphs

Except for the mask graph used as an anchor, all of the graphs were randomly ordered within each section. This was done to cancel out the order effect across the subjects (Shane:1992). Each package was stapled and assigned a test number. The packages, control and experimental, were randomly distributed to the subjects. After reading the instructions, the subjects were given a chance to ask questions about the conduct of the experiment. The experiment was strictly timed during execution to control for maturation. The subjects were given 15 seconds (except group four which got 30 seconds) to examine a graph and respond to the conclusion. The 15 second time limit was confirmed by the initial trial of the experiment to be of sufficient length to allow the subjects to accomplish the task. The forth group was given a 30 second time limit to provide additional verification that the 15 second limit was not inhibiting task performance. If the subjects were rushed with the 15 second limit, then increasing the limit to 30 seconds would be expected to obtain different results.

The experiment was administered to four different groups of subjects. The first group consisted of 32 (16-control, 16-experimental) graduate students attending the Air Force Institute of Technology (AFIT) master's degree program. The second group consisted of 43 (22-control, 21-experimental) Department of Defense (DoD) managers, both civilian and active duty military, attending Professional Continuing Education (PCE) courses at AFIT. The last two groups consist of undergraduate Business Management students at Ohio University. The third group contained 37 (18-control, 19-experimental) students, and the fourth group contained 35 (18-control, 17-experimental) students. The diversity of background and educational experience contained within these groups reduced the interaction of selection and the experimental factor and strengthens the generalizeability of the experimental results. For all of the groups, the experiment was conducted in a classroom environment familiar to the subjects. Table 6 shows the composition of the groups.

TABLE 6

COMPOSITION OF GROUPS

GROUP #	NUMBER OF CONTROL SUBJECTS	NUMBER OF EXPERIMENTAL SUBJECTS	TOTAL # IN GROUP	ACADEMIC SETTING	TIME FOR EACH GRAPH (SECONDS)
1	16	16	32	Graduate	15
2	22	21	43	Continuing Education	15
3	18	19	37	Undergraduate	15
4	18	17	35	Undergraduate	30

Statistical Analysis

This experiment was designed to test the hypothesis based on two independent samples. The variance of the population is unknown, therefore, the variances of the two samples are unknown and should be assumed to be unequal. The normality of the population distribution can be determined using the Wilk-Shapiro/Rankit Plot (Statistix User's Manual, 1991:242). A Wilk-Shapiro value greater than .9 may be considered normal (Reynolds, 1992). The size of each sample is greater than 30, so regardless of the normality of the underlying population distributions and variances, a two-sample t test can be used (Devore, 1991:338). If the population distribution is not normal, and can not be assumed to be so, a distribution free (nonparametric) test must be conducted. The Rank Sum test (Mann-Whitney U) is an appropriate test to conduct in this case (Devore, 1991:610). Because of the large sample sizes the more general z test could also be conducted (Devore, 1991:326). For this experiment the more restrictive tests, the two-sample t test and the Rank Sum test, are both conducted and the results compared to provide a safeguard against the effects of unknown population distribution and variance.

The level of significance (α) that was used in the evaluation of all statistical results was .05. This means if the P-value is less than or equal to α, Ho should be rejected with a confidence level of .95, and if the P-value is greater than α, Ho should not be rejected with the same level of confidence (Devore, 1991:315).

Research was conducted to determine the best way to measure the reliability of the experiment. Cronbach's Coefficient Alpha was investigated as a possible way of measuring the reliability. The major source of error within a test is due to the sampling of items. The error resulting from the sampling of items is entirely predictable from the average correlation. Consequently, coefficient alpha would be the correct measure of reliability for any type of item (Nunnally, 1978:226). Coefficient alpha was calculated on the data from the first experimental group. This calculation produced negative reliability values, indicating that coefficient alpha was not an appropriate reliability measure for this experiment (Shane, 1992). Since a measure of internal reliability could not be determined before the continuation of the experiment, the reliability was evaluated by comparing results between experimental groups after the fact. These results are discussed in Chapter IV.

In order to conduct the two-sample t and Rank Sum tests, the difference between a subject's responses for the pretest and posttest must be determined. This was done by totaling the pretest responses and the posttest responses, and then the difference (delta) between the pretest and posttest was determined by subtracting the total of the pretest scores from the total of the posttest scores. The mean delta values for each group (x_{bar} is the mean delta from the control group and y_{bar} is the mean delta from the experimental group) were used to determine the test statistic.

The following section describes the tests used to conduct the statistical analysis. After each statistical test is discussed, an example will be given showing the Statistix output of the test. Appendix C contains the exact steps that must be accomplished to obtain similar results using the Statistix statistical software program. Appendix D contains a description of the terms and abbreviations used in the Statistix output tables. Appendix D also describes the variables used in the analysis of the experimental results. Each example will use some of the dummy data displayed in Tables 7 and 8; Table 7 contains the dummy control group data, and Table 8 contains the dummy experimental group data. Because all of the data was placed in the same file, a "C" or an "E" was placed in front of each column heading to differentiate between the variables associated with the control and experimental groups, respectively. The variable transformations required to perform the statistical tests are contained in Table 9.

The test statistic for any t test is identified by a "t". This is done because "The test statistic is really the same here as in the large sample case [z], but is labeled T to emphasize that its null distribution is a t distribution with n-1 d.f. [degrees of freedom] rather than the standard normal (z) distribution (Devore, 1991:302). The formula for computing the two-sample t test is (Devore, 1991:339):

$$t = \frac{\overline{x}-\overline{y}-(\mu_1-\mu_2)}{S_p*\sqrt{\frac{1}{m}+\frac{1}{n}}} \qquad (2)$$

Where $x_{bar} - y_{bar}$ is the difference between the corresponding sample means, $\mu_1 - \mu_2$ is the difference between the assumed means of the population distributions, S_p is the square root of the pooled estimator sample variance, and m and n are the size of the two samples. Table 10 contains the Statistix output for the two-sample t test. This test, as well as all others within this chapter, were conducted using the dummy data contained in Tables 7 and 8, and the transformations in Table 9.

TABLE 7

DUMMY DATA FOR CONTROL GROUP

CASE	CGROUP	CP1	CP2	CC1	CC2	CAGE
1	1	2	7	1	8	1
2	1	3	5	1	9	2
3	1	2	4	2	7	3
4	2	1	3	2	8	2
5	2	1	2	1	9	3
6	2	3	4	3	7	1
7	3	2	3	1	7	4
8	3	1	6	3	8	3
9	3	3	3	2	9	3
10	3	1	1	3	9	2

TABLE 8

DUMMY DATA FOR EXPERIMENTAL GROUP

CASE	EGROUP	EP1	EP2	EE1	EE2	EAGE
1	1	9	4	1	7	1
2	1	8	3	2	5	4
3	1	7	2	3	6	3
4	2	9	3	2	8	3
5	2	8	2	2	9	2
6	2	7	1	1	6	3
7	2	6	3	2	9	1
8	3	8	1	1	8	2
9	3	9	3	2	6	4
10	3	8	2	3	7	2

TABLE 9

DUMMY DATA TRANSFORMATIONS

CASE	CPRETEST	CPOSTTEST	EPRETEST	EPOSTTEST	CDELTA	EDELTA	CDELTA1	EDELTA1
1	9	9	13	8	0	5	1	8
2	8	10	11	7	-2	4	2	6
3	6	9	9	9	-3	0	0	4
4	4	10	12	10	-6	2	-1	7
5	3	10	10	11	-7	-1	0	6
6	7	10	8	7	-3	1	0	6
7	5	8	9	11	-3	-2	1	4
8	7	11	9	9	-4	0	-2	7
9	6	11	12	8	-5	4	1	7
10	2	12	10	10	-10	0	-2	5

The P-value of .0001 indicates that there was a statistical difference between the control and experimental group responses.

In the Rank Sum test, all of the observations from each sample are combined, or pooled, into one sample with a size of m + n (m is the size of one sample, and n the size of the other). Then the observations are ordered (ranked) from smallest to largest, with the smallest receiving a rank of one. The test statistic for the Rank Sum test, "w", is computed based on the rank of the pooled observations. The formula for computing "w" is (Devore, 1991:612):

$$w = \sum_{i=1}^{m} r_i \tag{3}$$

Where r_i is the rank of the observation in the combined sample minus $\mu_1 - \mu_2$, and m is the number of observations in the smallest sample. Table 11 contains the Statistix output for the Rank Sum test. Once again, the P-value of .0001 suggests that there was a statistical difference between the control and experimental group responses. The result of this non-parametric test confirmed the result of the two-sample t test.

The final statistical procedure that was used is ANOVA. The term ANOVA, refers broadly to a collection of experimental situations and statistical procedures for the analysis of quantitative responses from experimental units. ANOVA involves the analysis of either data sampled from more than two numerical populations (distributions) or data from experiments in which more than two treatments have been used (Devore, 1991:371).

In this experiment ANOVA was conducted on both types of data. In the first case, ANOVA was used to determine if there was any statistical difference between the control and experimental group responses to the individual graphs, and to determine if there was any difference between the responses for line and bar graphs. This data were from two different distributions and were in a tabular format. ANOVA was also used to determine if there was any statistical difference in the responses between uniquely defined groups. This involved the analysis

TABLE 10

TWO SAMPLE T TESTS FOR CDELTA VS EDELTA

VARIABLE	MEAN	SAMPLE SIZE	S.D.	S.E.
CDELTA	-4.300	10	2.830	8.950E-01
EDELTA	1.300	10	2.359	7.461E-01

	T	DF	P
EQUAL VARIANCES	-4.81	18	0.0001
UNEQUAL VARIANCES	-4.81	17.4	0.0002

	F	NUM DF	DEN DF	P
TESTS FOR EQUALITY OF VARIANCES	1.44	9	9	0.2982

CASES INCLUDED 20   MISSING CASES 0

TABLE 11

RANK SUM TWO SAMPLE (MANN-WHITNEY) TEST FOR CDELTA VS EDELTA

VARIABLE	RANK SUM	SAMPLE SIZE	U STAT	AVERAGE RANK
CDELTA	59.00	10	4.000	5.9
EDELTA	151.0	10	96.00	15.1
TOTAL	210.0	20		

EXACT PROBABILITY OF A RESULT AS OR MORE EXTREME
THAN THE OBSERVED RANKS (1 TAILED P VALUE)   0.0001

NORMAL APPROXIMATION WITH CONTINUITY CORRECTION   3.439
TWO TAILED P VALUE FOR NORMAL APPROXIMATION   0.0006

TOTAL NUMBER OF VALUES WHICH WERE TIED   11
MAX. DIFF. ALLOWED BETWEEN TIES   1.0E-0005

CASES INCLUDED 20   MISSING CASES 0

of the demographic data with the various demographic factors considered to be treatments. This data was in a categorical format. The test statistic for ANOVA is "F", and the formula for computing "F" is:

$$F = \frac{MSTr}{MSE} \tag{4}$$

The formulas for computing MSTr and MSE are:

$$MSTr = \frac{J}{I-1}\sum_{i=1}^{I} (\bar{X}_{i}. - \bar{X}..)^2 \tag{5}$$

$$MSE = \frac{S_1^2 + S_2^2 + \ldots + S_I^2}{I} \tag{6}$$

Where MSTr is the unbiased estimator of σ^2, and MSE is an unbiased estimator of the common variance σ^2 (Devore, 1991:375). The reader should refer to a statistics text book for further explanation of MSTr and MSE. Table 12 contains the Statistix output of an ANOVA for responses to an individual graph. Table 13 contains the Statistix output of an ANOVA conducted on the control group demographic factor age. The P-value of .0000 in Table 12 indicates that there was a statistical difference between the control and experimental group responses for this particular graph. The P-value of .3910 in Table 13 indicates that there was no statistical difference between the responses of members of the control group when the subjects are segregated on the basis of their age.

Summary

The experiment followed the pretest-posttest control group design to determine if a scale break on the dependent (vertical) axis affects decision makers. The graphs were designed using Quattro Pro to incorporate specific characteristics. A two-sample t test and a Rank Sum test were conducted to test the hypothesis. ANOVA was used to determine if there was any significance between demographic factors. Chapter IV, Analysis and Findings, contains the results of the experiment.

TABLE 12

ONE WAY ANOVA FOR: CDELTA1 EDELTA1

SOURCE	DF	SS	MS	F	P
BETWEEN	1	180.0	180.0	101.25	0.0000
WITHIN	18	32.00	1.778		
TOTAL	19	212.0			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	0.00	1	1.0000

COCHRAN'S Q 0.5000
LARGEST VAR / SMALLEST VAR 1.000

COMPONENT OF VARIANCE FOR BETWEEN GROUPS 17.82
EFFECTIVE CELL SIZE 10.0

VARIABLE	MEAN	SAMPLE SIZE	GROUP STD DEV
CDELTA1	0.000	10	1.333
EDELTA1	6.000	10	1.333
TOTAL	3.000	20	1.333

CASES INCLUDED 20 MISSING CASES 0

TABLE 13

ONE WAY ANOVA FOR CDELTA = CAGE

SOURCE	DF	SS	MS	F	P
BETWEEN	3	26.85	8.950	1.19	0.3910
WITHIN	6	45.25	7.542		
TOTAL	9	72.10			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	1.46	2	0.4826

COCHRAN'S Q 0.6833
LARGEST VAR / SMALLEST VAR 5.486

COMPONENT OF VARIANCE FOR BETWEEN GROUPS 5.951E-01
EFFECTIVE CELL SIZE 2.4

CAGE	MEAN	SAMPLE SIZE	GROUP STD DEV
1	-1.500	2	2.121
2	-6.000	3	4.000
3	-4.750	4	1.708
4	-3.000	1	M
TOTAL	-4.300	10	2.746

CASES INCLUDED 10 MISSING CASES 0

IV. Analysis and Findings

Experimental Results

Appendix E contains the experiment results. Table 34A contains the control group responses and Table 34B contains the associated demographic data. Tables 35A and 35B contain the experimental group data. Appendix D contains a description of the terms and abreviations used in the Statistix output tables. Appendix D also describes the variables used in the experiment. Because the experimental data was collected from four unique groups (group composition is discussed in Chapter III.), the first step in data analysis was to conduct an ANOVA to determine if the results were statistically compatible between groups. This was done by using the individual designations in the Group column as categorical values and the data in the Delta column as the dependent variable under analysis. The results of this test are contained in Table 14 for the control group and Table 15 for the experimental group. The P-values of .5311 for the control group and .3387 for the experimental group indicated that there was no statistical difference between the mean responses for each of the groups, therefore, the responses from all of the groups were used in the analysis of the data.

Descriptive statistics were calculated to gain an understanding of the general nature of the data collected. The results for the control group are contained in Table 16, and the results for the experimental group in Table 17. An analysis of these results showed that the responses for each of the graphs, with the exception of graph C1/E1, were in the direction intended by the researchers (graph features are shown in Tables 3 and 4). The response for graph C1/E1 was in the direction of "agree" versus "disagree." A closer examination of this graph showed that the observed response was reasonable. The C1 graph depicts annual sales and shows an increase of five million dollars from 1954 to 1956, an increase of twenty percent. It is reasonable to

TABLE 14

ONE WAY ANOVA FOR CONTROL DELTA (CDELTA) = CONTROL GROUP (CGROUP)

SOURCE	DF	SS	MS	F	P
BETWEEN	3	104.8	34.93	0.75	0.5311
WITHIN	70	3.275E+03	46.79		
TOTAL	73	3.380E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	2.22	3	0.5284

COCHRAN'S Q 0.3103
LARGEST VAR / SMALLEST VAR 1.943
COMPONENT OF VARIANCE FOR BETWEEN GROUPS -6.440E-01
EFFECTIVE CELL SIZE 18.4

CGROUP	MEAN	SAMPLE SIZE	GROUP STD DEV
1	3.125	16	7.667
2	3.045	22	6.565
3	5.611	18	5.500
4	2.500	18	7.571
TOTAL	3.554	74	6.841

TABLE 15

ONE WAY ANOVA FOR EXPERIMENTAL DELTA (EDELTA) =

EXPERIMENTAL GROUP (EGROUP)

SOURCE	DF	SS	MS	F	P
BETWEEN	3	208.9	69.62	1.14	0.3387
WITHIN	69	4.208E+03	60.98		
TOTAL	72	4.417E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	1.74	3	0.6289

COCHRAN'S Q 0.3194
LARGEST VAR / SMALLEST VAR 1.854
COMPONENT OF VARIANCE FOR BETWEEN GROUPS 4.751E-01
EFFECTIVE CELL SIZE 18.2

EGROUP	MEAN	SAMPLE SIZE	GROUP STD DEV
1	6.250E-01	16	7.982
2	-1.857	21	8.799
3	2.105	19	6.463
4	2.118	17	7.713
TOTAL	6.438E-01	73	7.809

consider a twenty percent increase in sales to be a "significant" increase.

TABLE 16

DESCRIPTIVE STATISTICS FOR CONTROL GROUP RESPONSES

VARIABLE	CASES	MEAN	S.D.	SE(MEAN)	MINIMUM	MAXIMUM
P1	74	2.919	1.841	2.141E-01	1.000	9.000
P2	74	8.216	8.955E-01	1.041E-01	5.000	9.000
P3	74	5.216	3.189	3.708E-01	1.000	9.000
P4	74	3.311	2.548	2.962E-01	1.000	9.000
P5	74	5.703	1.957	2.274E-01	1.000	9.000
P6	74	4.122	2.027	2.356E-01	1.000	8.000
C1	74	6.189	2.092	2.431E-01	1.000	9.000
C2	74	6.608	1.662	1.932E-01	3.000	9.000
C3	74	1.527	7.977E-01	9.272E-02	1.000	5.000
C4	74	6.811	1.741	2.024E-01	1.000	9.000
C5	74	5.176	2.290	2.662E-01	1.000	9.000
C6	74	6.730	1.746	2.030E-01	1.000	9.000

TABLE 17

DESCRIPTIVE STATISTICS FOR EXPERIMENTAL GROUP RESPONSES

VARIABLE	CASES	MEAN	S.D.	SE(MEAN)	MINIMUM	MAXIMUM
P1	73	2.753	1.665	1.949E-01	1.000	9.000
P2	73	8.041	1.399	1.637E-01	1.000	9.000
P3	73	5.877	3.113	3.644E-01	1.000	9.000
P4	73	3.247	2.272	2.659E-01	1.000	9.000
P5	73	5.877	2.153	2.520E-01	1.000	9.000
P6	73	4.247	1.998	2.339E-01	1.000	9.000
E1	73	6.890	2.227	2.607E-01	1.000	9.000
E2	73	7.000	2.048	2.397E-01	1.000	9.000
E3	73	1.644	9.032E-01	1.057E-01	1.000	5.000
E4	73	3.699	2.259	2.644E-01	1.000	9.000
E5	73	5.342	2.237	2.619E-01	1.000	9.000
E6	73	6.110	2.059	2.409E-01	1.000	9.000

Before testing the hypotheses, the normality of the responses (Delta) was evaluated using the Wilk-Shapiro/Rankit Plot. Figure 11 contains the Wilk-Shapiro plot for the control group, and Figure 12 contains the plot for the experimental group. The results indicate that the observations for both the control group (.9831) and the experimental group (.9425) are reasonably normally distributed. Histograms of the response values shown in Figure 13 for the control group and Figure 14 for the experimental group, approximate the bell shaped curve characteristic associated with the normal distribution.

```
              RANKITS VS CONTROL DELTA
RANKITS
   3 +
     |
     |                                      ++
     |                                 +  +
     |                             +2 +
   1 +                           44
     |                        2 6+
     |                    23+24
     |                  534
     |                27
  -1 +            +34
     |          2++
     |    +    +
     | 2
     |
  -3 +
     -+---------+---------+---------+---------+-
     -15       -6          3        12        21
                     CONTROL DELTA
       APPROX. WILK-SHAPIRO 0.9831   74 CASES PLOTTED
```

Figure 11. Normality Plot for Control Group Responses

```
              RANKITS VS EXPERIMENTAL DELTA
RANKITS
   3 +
     I
     I                                2
     I                                2
     I                              4
   1 +                             34
     I                           28
     I                          93
     I                        722
     I                       37
  -1 +                     +5+
     I                    3+
     I                 ++
     I          +     +
     I
  -3 +
     -+---------+---------+---------+---------+-
     -50       -30       -10        10        30
                    EXPERIMENTAL DELTA
       APPROX. WILK-SHAPIRO 0.9425   73 CASES PLOTTED
```

Figure 12. Normality Plot for Experimental Group Responses

```
FREQUENCY                       HISTOGRAM OF CONTROL DELTA
    |
 18 +
    |                        ******
    |                  ************
 15 +                  ******************
    |                  ******************
    |                  ******************
 12 +                  ************************
    |                  ************************
    |                  ************************
  9 +                  ************************
    |                  ************************
    |                  ************************
  6 +            ******************************
    |            ******************************
    |            ******************************
  3 +            ******************************      ******
    |******      ******************************************
    |******************************************************
  0 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
  -15   -11    -7    -3     1     5     9    13    17    21
```

Figure 13. Histogram of Control Group Delta

```
FREQUENCY              HISTOGRAM OF EXPERIMENTAL DELTA
    |
 18 +
    |
    |                                   *****
 15 +                                   **********
    |                                   **********
    |                                   **********
 12 +                              ********************
    |                              ********************
    |                              ********************
  9 +                              ********************
    |                              ********************
    |                              *************************
  6 +                              *************************
    |                         ******************************
    |                         ******************************
  3 +                         ******************************
    |               ****************************************
    |*****          *********************************************
  0 +----+----+----+----+----+----+----+----+----+----+----+----+
  -30  -26  -22  -18  -14  -10   -6   -2    2    6   10   14   18
```

Figure 14. Histogram of Experimental Group Delta

The next step was to evaluate the basic experimental hypothesis, "Are graphs with scale breaks on the dependent axis interpreted differently from graphs without a scale break." The tests used in this evaluation were the two-sample t test and the Rank Sum test. As identified in Chapter III, the hypothesis can be written as:

$$Ho: \quad (O_2 - O_1) - (O_4 - O_3) = 0$$

$$Ha: \quad (O_2 - O_1) - (O_4 - O_3) \neq 0$$

These tests compared the control group Delta values with the experimental group Delta values. The results of the t test are contained in Table 18, and the results of the Rank Sum test are contained in Table 19. The one-tailed P-value of .0174 (two-tailed value is .0348) obtained by the t test closely corresponds with the two-tailed P-value of .0333 produced by the Rank Sum test. The results of both tests indicated that the null hypothesis should be rejected in favor of the alternative hypothesis with a reasonable level of confidence. So, it has been empirically demonstrated that there was a difference in interpretation between graphs with and without scale breaks.

An ANOVA was conducted to determine the impact of each individual experimental graph on the difference between the pretest and posttest responses of the control group and the experimental group for each posttest pair (C1, E1). This difference was determined by subtracting the posttest response from the appropriate pretest response after the responses were standardized to an agree response. Table 20 shows the matching of the posttest and pretest graphs. "Inverse", as used in the table, means that the original values was subtracted from 10 so that all responses are based on an agree indication.

A similar ANOVA was conducted between the posttest graphs that were designed to have dramatic lie factors (E1 E4 E5) and the graphs designed to have nondramatic lie factors (E2 E3 E6). The difference between the pretest and posttest line graph responses, and the

TABLE 18

TWO SAMPLE T TESTS FOR CONTROL DELTA (CDELTA) VS

EXPERIMENTAL DELTA (EDELTA)

VARIABLE	MEAN	SAMPLE SIZE	S.D.	S.E.
CDELTA	3.554	74	6.805	7.910E-01
EDELTA	6.438E-01	73	7.832	9.167E-01

	T	DF	P
EQUAL VARIANCES	2.41	145	0.0174
UNEQUAL VARIANCES	2.40	141.7	0.0167

	F	NUM DF	DEN DF	P
TESTS FOR EQUALITY OF VARIANCES	1.32	72	73	0.1166

TABLE 19

RANK SUM TWO SAMPLE (MANN-WHITNEY) TEST FOR CONTROL DELTA

(CDELTA) VS EXPERIMENTAL DELTA (EDELTA)

VARIABLE	RANK SUM	SAMPLE SIZE	U STAT	AVERAGE RANK
CDELTA	6.026E+03	74	3.251E+03	81.4
EDELTA	4.852E+03	73	2.151E+03	66.5
TOTAL	1.088E+04	147		

NORMAL APPROXIMATION WITH CONTINUITY CORRECTION 2.129
TWO TAILED P VALUE FOR NORMAL APPROXIMATION 0.0333

TOTAL NUMBER OF VALUES WHICH WERE TIED 138
MAX. DIFF. ALLOWED BETWEEN TIES 1.0E-0005

TABLE 20

PRETEST POSTTEST GRAPH MATCHING

POSTTEST GRAPH	PRETEST GRAPH
1	2
2	6 (Inverse)
3 (Inverse)	3
4	5
5	4 (Inverse)
6	1 (Inverse)

difference between the bar graph responses was also analyzed using ANOVA. The results of these ANOVAs are summarized in Table 21, and the detailed results of the ANOVAs are contained in Appendix F, Tables 36 to 47. An asteristic indicates that the graph with the greatest distortion was excluded from the evaluated group. These results indicate that graphs 1 and 4 show significance between the pretest and posttest responses. And a grouping of all the graphs that are considered to have a dramatic lie factor, do, in fact, have a statistically significant difference between the control and experimental responses. The graphs that are considered to have a nondramatic lie factor do not have a statistically significant difference between responses, either individually or as a group. This result, which implies a correlation between the lie factor and the interpretation of a graph, disagrees with a conclusion made by Kern that "as the magnitude of the lie factor increases, the degree to which a decision maker is mislead may or may not increase" (Kern 1991:40). This conflict could be the result of differences in experimental design between the two experiments. The graphs in Kern's experiment were exactly matched between pretest and posttest, while in this experiment the graphs were designed so that the tasks for the pretest graphs and the posttest graphs were generally similar, but not exactly the same.

An analysis of the combined responses to line graphs showed a difference in interpretation between line graphs with and without scale breaks. However, a difference in interpretation was not evident for the combined bar graph responses. To reduce the influence produced by the graphs with the greatest amount of distortion, line and bar graph responses were also evaluated with the graph with the largest lie factor removed from each group. This action had a major impact on the analysis in absolute terms. The p-value for each group increased by approximately .4, however, the difference between the line and bar graph groups remained constant. These results tend to indicate that the interpretation of line graphs are affected more strongly by scale break than bar graphs.

An ANOVA was also conducted on the demographic data to determine if there were any personal characteristics that lead to differing results between respondents. The complete results of the ANOVA are contained in Appendix F, Tables 48 to 61. A summary of the results is contained in Table 22. The asteristic indicates a significant difference between the subjects' responses when they were segregated based on the particular demographic factor.

This demographic analysis indicated that there was a difference between the experimental group responses based on sex and professional experience. Men tended to respond differently to the broken scale then women did, and subjects with experience in different fields responded differently to the broken scale. There was no difference between these factors for the control group. The "graph use" factor also indicated a significant difference between responses. Those subjects that use graphs more frequently responded differently from those who do not use graphs very often.

<u>Summary</u>

Analysis of the data indicates that there was a difference in the way subjects responded to graphs with a scale break and graphs without a scale break. Only two graphs, numbers 1 and 4, when looked at independently showed any significant difference in the responses between the control and experimental groups. The group of graphs intended to have a nondramatic lie factor had no difference in the responses, while the group of graphs intended to have a dramatic lie factor had a difference in the responses. The subject's sex and professional experience showed a difference between the responses of the experimental group. Familiarity with graphs and their use showed a difference between responses for both groups.

TABLE 21

SUMMARY OF ANOVA RESULTS

COMPARISON FACTORS	LIE FACTOR	P-VALUE
C4 vs E4	36.8	.0000
C1 vs E1	11.8	.0303
C5 vs E5	9.2	.8481
C2 vs E2	6.3	.2585
C3 vs E3	2.7	.1414
C6 vs E6	2.7	.0965
C2 C3 C6 vs E2 E3 E6 (Nondramatic)	< 8	.2349
C1 C4 C5 vs E1 E4 E5 (Dramatic)	> 8	.0042
C3 C4 C5 vs E3 E4 E5 (Line Graphs)	N/A	.0000
C3 C5 vs E3 E5 * (Line Graphs)	N/A	.3866 *
C1 C2 C6 vs E1 E2 E6 (Bar Graphs)	N/A	.3002
C2 C6 vs E2 E6 * (Bar Graphs)	N/A	.7066 *

TABLE 22

ANOVA OF DEMOGRAPHIC DATA

DEMOGRAPHIC FACTOR	P-VALUE FOR CONTROL GROUP	P-VALUE FOR EXPERIMENTAL GROUP
Sex	.8483	.0336 *
Age	.6277	.1359
Education Level	.9229	.7547
Professional Experience	.9029	.0001 *
Employment	.1346	.4898
Graph Use	.0447 *	.0412 *
Graph Construction	.8447	.9429

V. Conclusion

Many decisions are made based on information presented in a graphical format. The possibility exists for this information to be misrepresented. Bad or incorrect decisions may be made based on this misrepresented data. The widespread use of computers, and the variety of computer graphics programs available have made it easier to create high-quality graphs. Because of the variety of features in the graphics programs, the uninformed user can create poorly constructed graphs that misrepresent the data and mislead decision makers.

Summary of Results

A review of the literature provided answers to three of the investigative questions posed by this research. An analysis of current trends in business graphics revealed that there are a wide variety of computer graphics programs available. The impetus for the creation of such a large selection of graphics applications has been the wide spread use of personal computers and the perceived benefits of graphical representation. In an effort to satisfy market demand, the different programs offer a diversity of graphic construction features and tremendous flexibility. However, this flexibility has provided the potential for graphs to be formatted incorrectly by the uninformed or deception-minded graphics producer. If a graph is drawn with an improper format, the effectiveness of any decisions based on that graph can be diluted.

One of the major formatting issues which can affect the effectiveness of a graph is the construction of the dependent axis scaling. It has been pointed out that in certain decision making situations the need for graph resolution takes precedent over the need to construct a continuous scale portraying the full range of data values (Cleveland, 1985:79). Two methods exist for manipulating the axis scaling to achieve greater resolution: start the scale at a non-zero point, or break the scale and eliminate some intermediate range of values. Most current software packages allow for the first option but not the second. Scale breaks can be constructed, however, if the graphics application has a graph annotator or drawing feature.

The issue of whether or not scale breaks are appropriate gets a mixed review in the literature. The authors of graph construction criteria and style guides are fairly evenly split in their opinion of this subject. Unfortunately, sources on neither side of the issue have the weight of empirical evidence to justify their position.

An experiment was conducted to determine if decision makers interpreted data presented in a graph with a scale break on the dependent (vertical) axis differently than graphs without a scale break. One hundred forty-seven subjects were randomly divided into two groups, control and experimental, and administered a pen and paper experiment. The control group reviewed graphs that followed criteria for high-integrity graph construction. The experimental group reviewed similar graphs except for posttest graphs which had a scale break on the dependent axis.

Two-sample t and Rank Sum tests were conducted on the combined experimental results to test the overall null hypothesis that scale breaks do not affect the interpretation of graphs. Both tests provided the same result; graphs with a scale break were interpreted differently from those without a scale break. This finding should not lead to the conclusion that a scale break should never be used. Decision makers should be aware of the effect scale break has on interpretation and allow the situation to dictate the appropriateness of its use. The type of task being undertaken, and the resolution required to make an accurate decision should be the primary factors considered.

Graph responses were analyzed individually and in groups to determine the correlation between graph interpretation and the amount of visual distortion. When analyzed individually, it was determined that there was no significant difference between responses for graphs with a small lie factor (<8), but a difference was indicated for two of the three graphs with a larger lie factor (>8). Graph 5 was the only large lie factor graph with no significant difference. A closer review of the task associated with this graph revealed that the information conveyed by the graph was not affected by the scale break. The task was to determine if the data was stabilizing, and this trend was evident in both the pretest and posttest graphs.

The combined analysis of the graphs with dramatic and nondramatic lie factors produced the expected results; graphs with dramatic lie factors showed a significant difference and graphs with nondramatic lie factors did not. Therefore, it can be concluded that the amount of visual distortion present in a graph is correlated with graphic interpretation. This result is at odds with the findings of Kern's research.

The individual graph responses were also grouped to allow an analysis of the effect of scale break on line and bar graphs. This analysis indicated that line graphs are affected more than bar graphs. The ANOVA of the three line graphs produced a p-value of .0000, while the three bar graphs produced a p-value of .3002. The most distorted line graphs had a lie factor of 36.8, while the most distorted bar graph had a lie factor of only 11.8. To remove the possible bias induced by the 36.8 lie factor graph, the most dramatic graph was eliminated and the analysis was reaccomplished. This analysis produced a p-value of .3866 for the line graphs, and a p-value of .7066 for the bar graphs. Since the difference between graph types for both of these groupings was approximately the same (.4), the researchers concluded that a scale break has a differing influence on line and bar graphs. It is possible that this effect could be attributed to an inappropriate experimental design, or to a difference in task analysis.

The demographic factors associated with the individual subjects were analyzed to determine if there were any factors that contributed to a difference in responses. The sex of the subject made a difference in the experimental group; males responded to the scale break differently then women. The professional background of the subject was also a factor that led to a difference in responses for the experimental group. The extent to which the subject had experience using graphs in decision making was correlated with a difference in response for both groups.

Recommendations for Future Research

Previous research showed that graphs constructed with a non-zero axis affected interpretation of the graphics. This experiment showed that graphs constructed with a scale break can also affect decision makers. Both of these techniques of modifying the dependent axis scale can be used to increase the resolution of data portrayed by a graphic representation. It is not known, however, which of these two techniques produces the most accurate interpretation or is preferred by decision makers. Future research should be conducted to determine which method of increasing graphic resolution is the most appropriate.

The research conducted by Kern failed to show a correlation between the magnitude of the lie factor and the degree to which a graph was misleading (Kern. 1991:40). This research indicated that a correlation between the lie factor and graph interpretation accuracy may exist. Future research could investigate, in greater detail, the relationship between the lie factor and the interpretation of graphs. The ultimate objective would be the identification of a visual distortion threshold to prevent the construction of misleading graphs.

The authors rejected the computer format for the execution of the experiment because of the possibility of confounding variables associated with the computer format. Future research could be conducted to identify the variables associated with using a computer format and how to control for these variables.

One of the more difficult aspects of preparing this experiment was the effort required to produce high quality representations of scale breaks within the Quattro Pro graph annotator. It was obvious that the features of the annotator were not designed with such an application in mind. This short coming points out the possibility that the ability to accurately represent scale breaks and other recommended graph style techniques may not exist in popular software packages. It would be valuable to have a survey of current spreadsheet, graphics, and presentation applications to ascertain their ability to adhere to proven graph construction criteria. This would allow potential users of these packages to evaluate the software programs based on their capacity to produce high integrity graphics.

In addition, this and previous research investigated many of the existing standards for developing high integrity graphs. However, the list of criteria and guides has not been exhausted. Any future effort which would continue this work would help to build a more refined list of usable graphics standards.

Recommendations

As research into the subject of graph construction criteria is accomplished, the results should be consolidated into a single source of usable standards. The current situation forces a graphic artist concerned with standard compliance to review a variety of sources (provided he is aware of the existence of the multiple sources). He must then exercise personal judgment to determine which of the conflicting criteria is most appropriate. The consequences of faulty decisions made as a result of improperly formatted graphics are too great to allow personal preference to be the ultimate criteria for presentation construction. Formatting standards should be scientifically demonstrated and easily available.

It was assumed by the authors that graphs constructed using the standard settings of the graphics software and a full range of values would produce graphs with lie factors very close to one. After the graphs were constructed and the lie factors determined, the authors discovered that the graphs, except the posttest graphs for the experimental group, had a lie factors ranging from .903 to 1.13. This discovery points out how easy it is to induce distortion in graphic representations, either intentionally or unintentionally. It would be beneficial for software manufacturers to incorporate a safety mechanism in their graphics applications that would provide an indication of the visual distortion present in a graph. This feature could either be an indication of the lie factor value of the graph or a warning that the graph is exceeding a predetermined distortion threshold.

Appendix A: Criteria and Style Guides for Construction of High Integrity Graphs

This appendix contains tables which provide an author cross-referenced matrix of graph construction techniques. An author's agreement with a particular criterion or guideline is indicated with an "X", while an author's disagreement with a particular criterion is indicated with an "O". The authors cross-referenced in the tables are as follows:

	Author	Year
1.	Tufte	1983
2.	Taylor	1983
3.	Larkin	1990
4.	Schmid and Schmid	1954
5.	Joint Committee on Standards for Graphic Representation	1915
6.	MacGregor	1979
7.	Steinbart	1986
8.	Johnson, Rice, and Roomich	1980
9.	Spear	1969
10.	Auger	1979
11.	Rogers	1961
12.	American Society of Mechanical Engineers	1979
13.	Lefferts	1981
14.	Cleveland	1985
15.	Schmid	1983

TABLE 23

CRITERIA FOR CONSTRUCTING HIGH INTEGRITY

GRAPHICS, CROSS REFERENCED BY AUTHOR

CRITERIA	AUTHORS: 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Charts with a arithmetic scale should begin at the zero base line.	X	X		X	X	X	X	X	X	X	X	X	X	O	X
2. Use multiple scales cautiously.		X		X								X			
3. The dependent axis should employ a simple arithmetic scale.						X		X				X			
4. Do not extend the scale much beyond the highest or lowest points on the graph.		X		X							X	X		X	
5. If multiple curves are shown, the same unit scale must be used.													X		
6. Use labels to reduce graphical distortion and ambiguity.	X		X	X					X			X			
7. Represent quantities by linear magnitudes as areas or volumes may be misinterpreted.	X		X	X	X	X			X	X					
8. For area graphs, the more irregular strata should be placed near the top.			X												
9. Time scale divisions must be equal.				X		X				X	X	X			
CRITERIA	AUTHORS: 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

TABLE 23 CONTINUED

CRITERIA	AUTHORS:														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
10. Keep charts simple to add to clarity.											X	X	X		
11. The general arrangement of a graph should be from left to right and bottom to top.			X		X										
CRITERIA	AUTHORS:														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

TABLE 24

GRAPHICS CONSTRUCTION STYLE GUIDELINES,

CROSS REFERENCED BY AUTHOR

CRITERIA	AUTHORS: 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Scale breaks should be used for false origins.				X	X							X			
2. Graphics must not quote data out of context.	X														
3. Oblong shaped grids are preferable to square grids.				X								X			
4. The zero line should be sharply distinguished.				X	X				X						
5. The curve lines should be distinguished from the grid ruling.					X				X		X	X			
6. If a diagram does not include data, it should accompany the chart in tabular form.				X							X	X			
7. When shading, shade from the zero line to the curve.									X						
8. Vertical or horizontal shadings not recommended.									X			X	X		
9. Patterned shadings should be of good contrast.				X		X				X	X	X			
CRITERIA	AUTHORS: 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

TABLE 24 CONTINUED

CRITERIA	AUTHORS: 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
10. For column charts, the columns should be the same width; spacing between equal to one-half the column width.						X			X						
11. Arrange columns systematically.									X		X				
12. When a large part of the grid is unnecessary, break the grid but retain the zero line.						X									
13. Eliminate all grid lines but those essential for easy reading.	X			X		X					X	X			
14. On multiple scale curve graphs, each curve should be the same width.						X					X	O			
15. If irregularities occur in the time sequence, include spaces for the missing periods.									X						
16. Avoid broken scales which give inaccurate impressions.				X						X		X		X	
CRITERIA	AUTHORS: 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

TABLE 24 CONTINUED

CRITERIA	AUTHORS: 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
17. Standardized units of monetary measurement are better than nominal units.	X														
18. For most line charts the maximum number of plotted lines should not exceed five; three or fewer is better.													X		
19. The simplest curve patterns are usually the most effective. A solid line is most useful.												X			
20. Keep charts as simple as possible to add to clarity.											X			X	
21. Do not overdo the number of tick marks.														X	
CRITERIA	AUTHORS: 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Appendix B: Experimental Instrument Contents

This appendix contains all the information contained in the experimental instrument used in this research. The information is presented in a manner structured to accentuate the relationships between the separate components of the instrument and between the control and experimental instruments. The cover sheet/instructions are shown first. The graphs used as masks are shown next. The pretest graphs follow the mask graphs. Each of the experimental instruments contained all of the mask and pretest graphs, but the order of presentation of the graphs was randomly arranged for each instrument. The final section of graphs contains both the control and experimental post-test graphs; each control and experimental graph pair are presented on the same page. Although these graphs are shown together on the same page, each group (control and experimental) only saw one of the paired graphs. The final item presented is the demographic questionnaire.

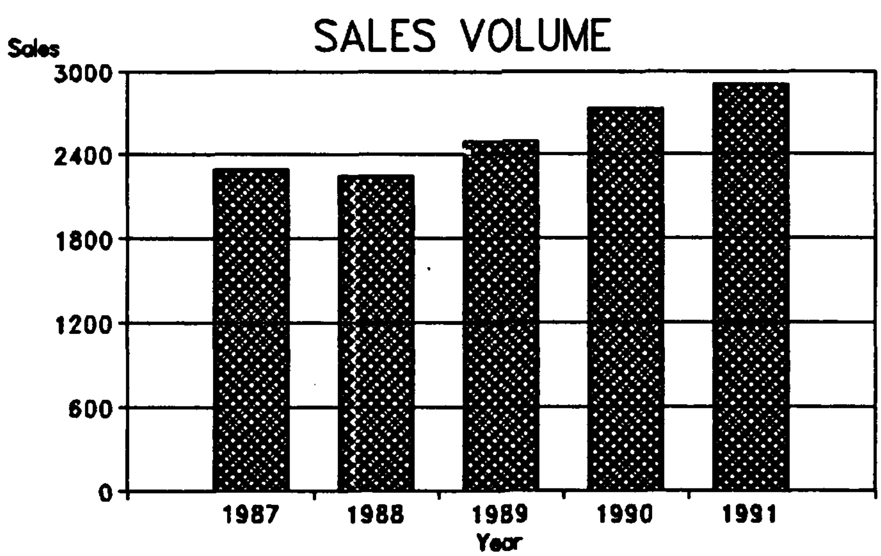
INSTRUCTIONS

Introduction: First of all, thank you for your participation.

This exercise is being conducted as part of an AFIT thesis research project. The results of the project may be used to improve the quality and content of decision support systems both in the Air Force and civilian communities. In this exercise you will view a sequence of 18 different graphs. You will then be asked to respond to a statement about the graph. Your response should be based on your impressions generated from viewing the graph - there are no right or wrong answers. For the purposes of this experiment any significant trend or change can be either a "significant increase" or a "significant decrease." Each graph should be analyzed independently of the others.

Example:

The following example illustrates the format of the graph, conclusion, and the response scale.



Conclusion: Sales volume remained fairly constant throughout the period.

Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree

(Turn to the next page)

Disclaimer:

The data represented in this experiment are fictional and developed purely for the purposes of this exercise.

Length:

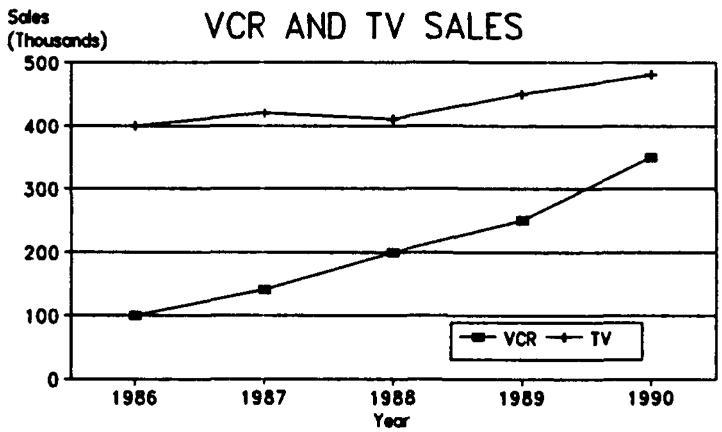
This exercise will be timed. You will be given 15 seconds per page to view the graph and mark your response. The monitor will tell you when to turn to the next page.

The total time required for this exercise should not exceed 15 minutes.

When to Start:

PLEASE WAIT FOR THE MONITOR'S SIGNAL TO BEGIN.

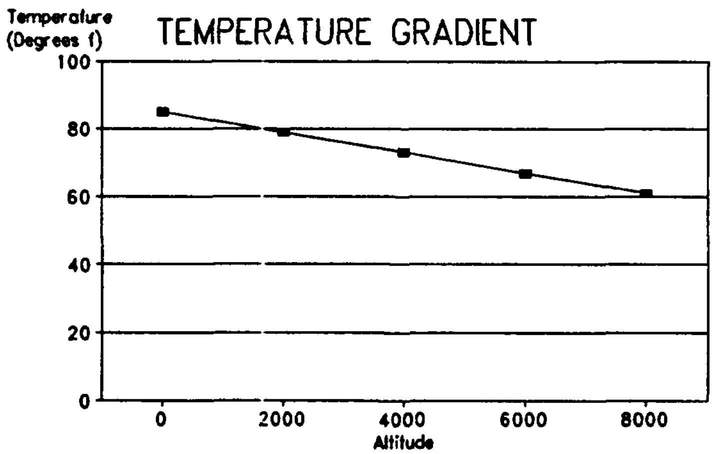
Thank you again for your cooperation!



Conclusion: TV sales remained relatively constant, when compared to VCR sales.

Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree

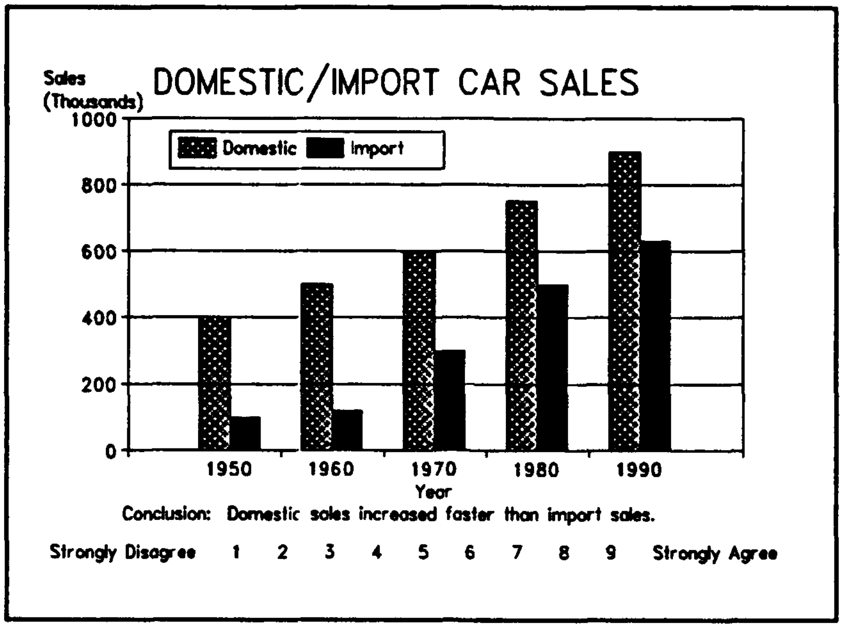
Mask Graph 1



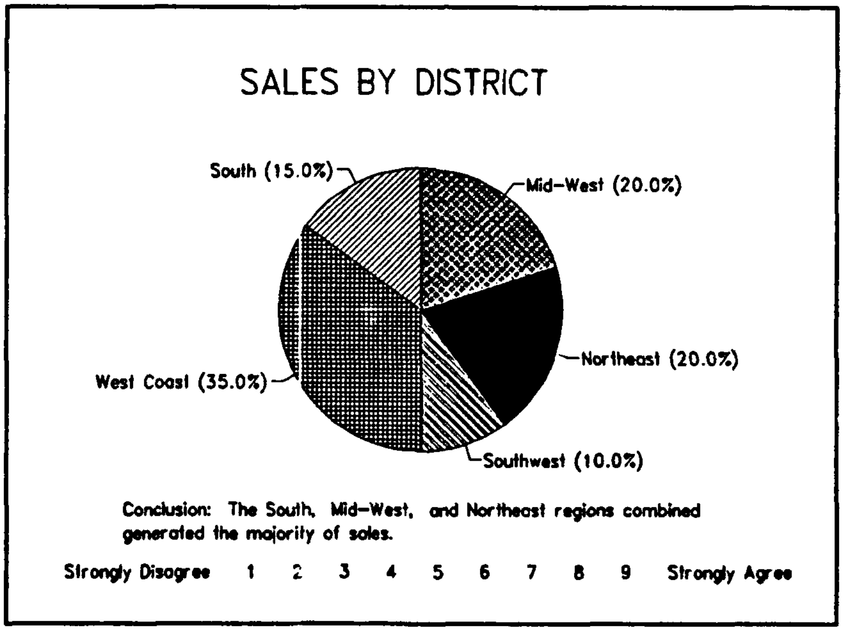
Conclusion: Temperature falls significantly as altitude increases from sea level to 8, 000 feet.

Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree

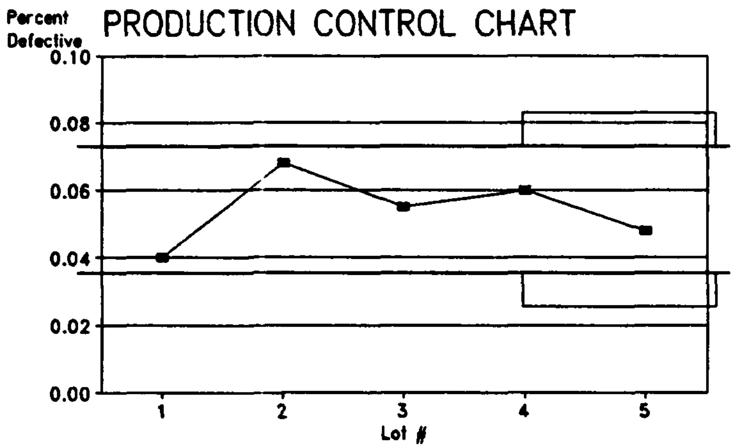
Mask Graph 2



Mask Graph 3



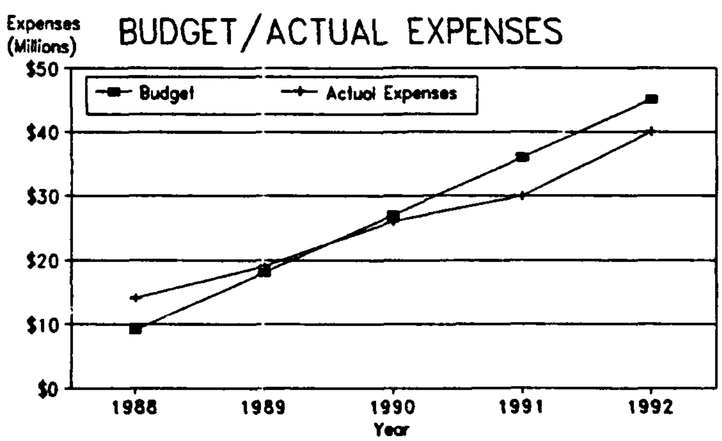
Mask Graph 4



Conclusion: The production process is producing output outside the specified tolerance.

Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree

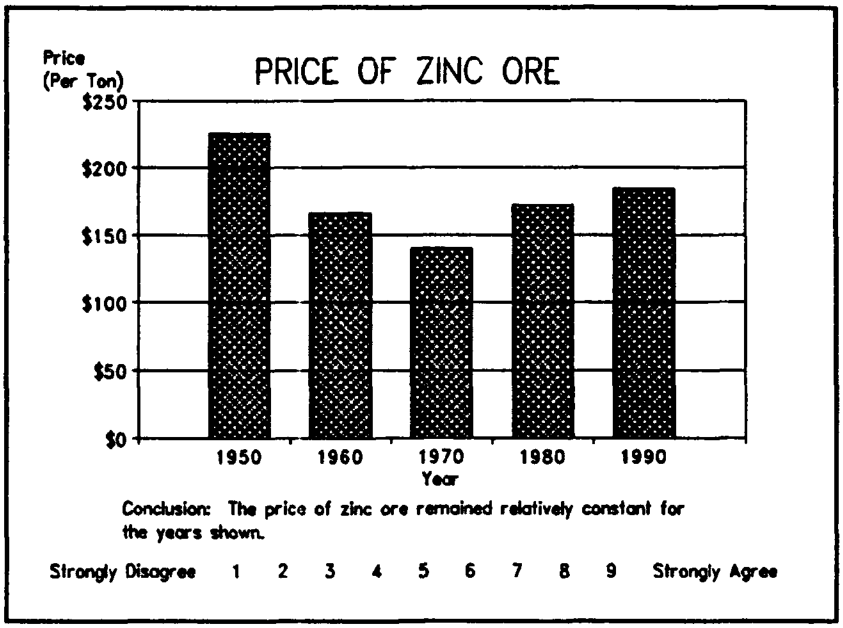
Mask Graph 5



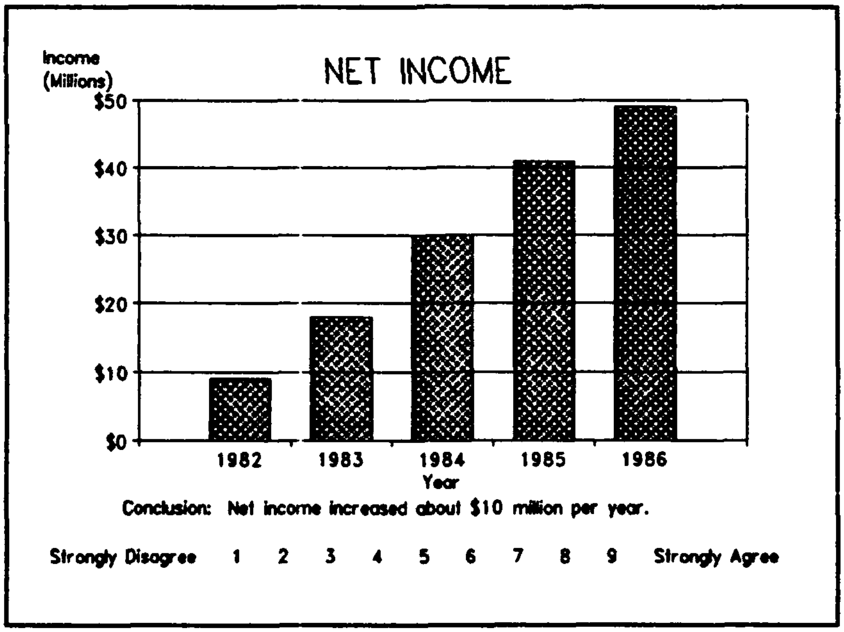
Conclusion: Actual expenses remained consistently under the planned budget.

Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree

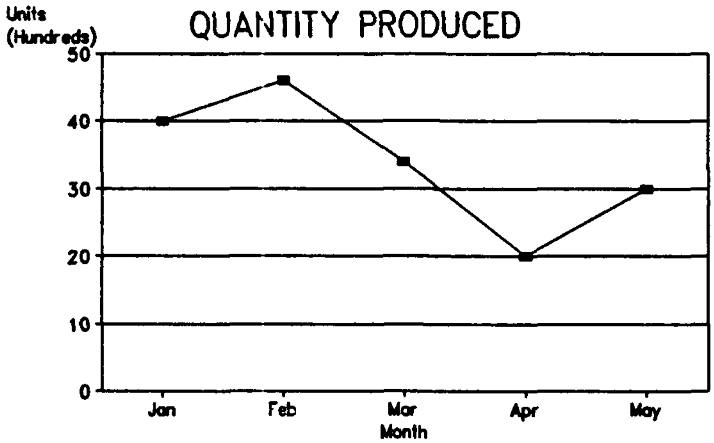
Mask Graph 6



Pretest Graph 1



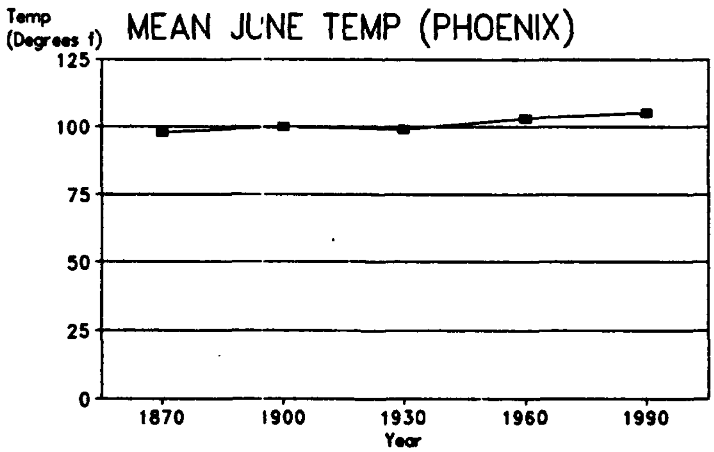
Pretest Graph 2



Conclusion: The quantity produced declined significantly during April.

Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree

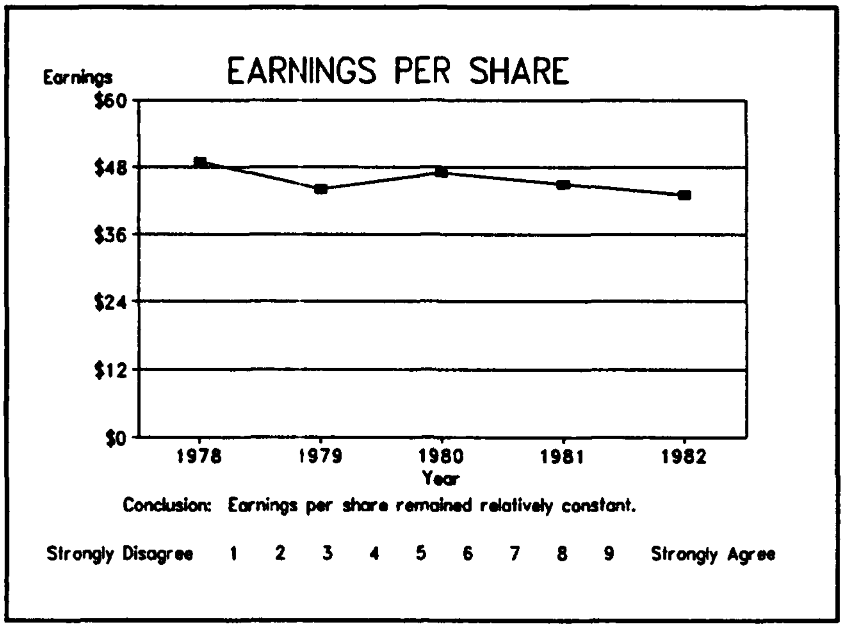
Pretest Graph 3



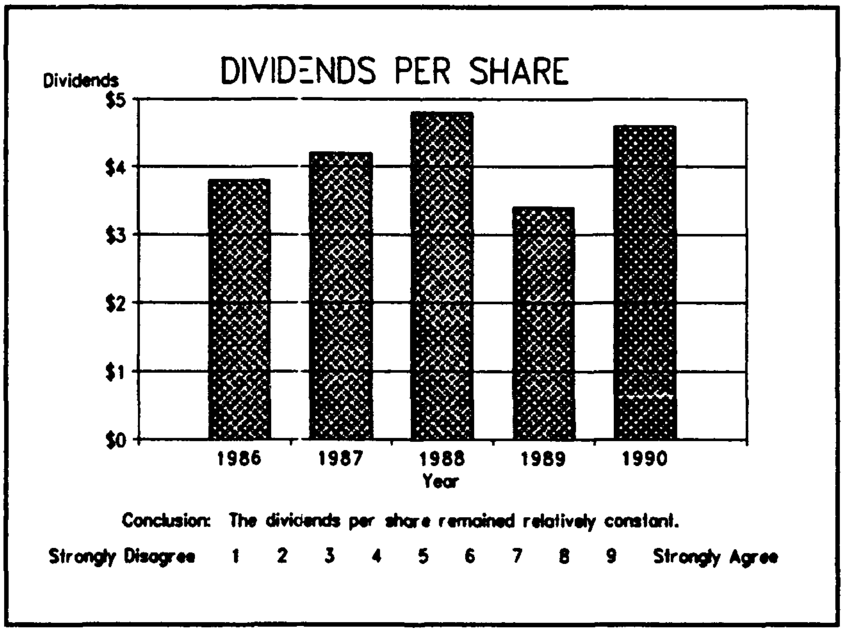
Conclusion: The mean June temperature in Phoenix has increased significantly over the last 80 years.

Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree

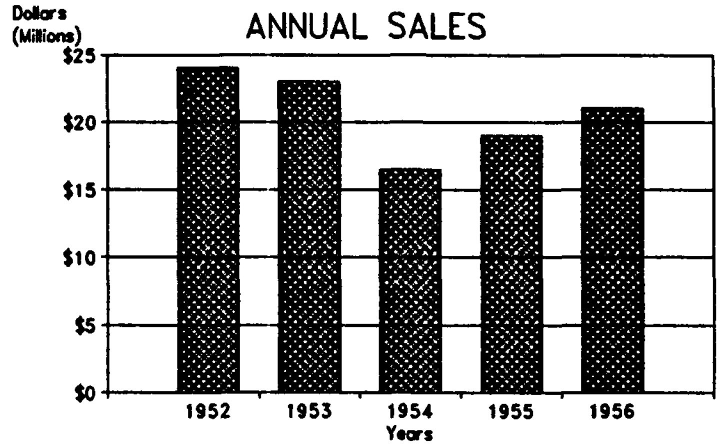
Pretest Graph 4



Pretest Graph 5



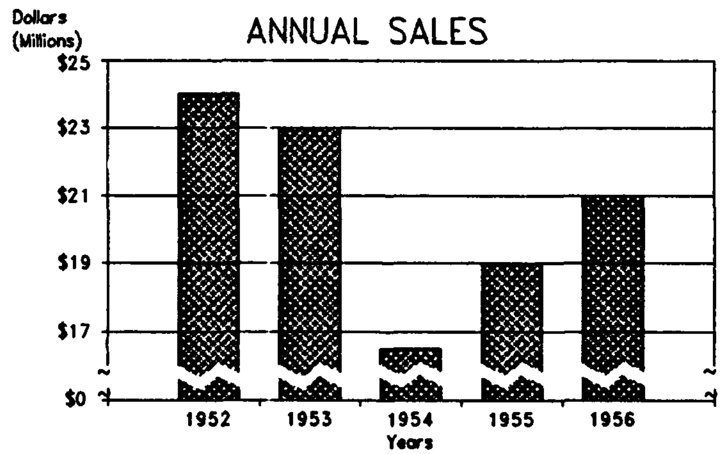
Pretest Graph 6



Conclusion: Annual sales increased significantly from 1954 to 1956.

Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree

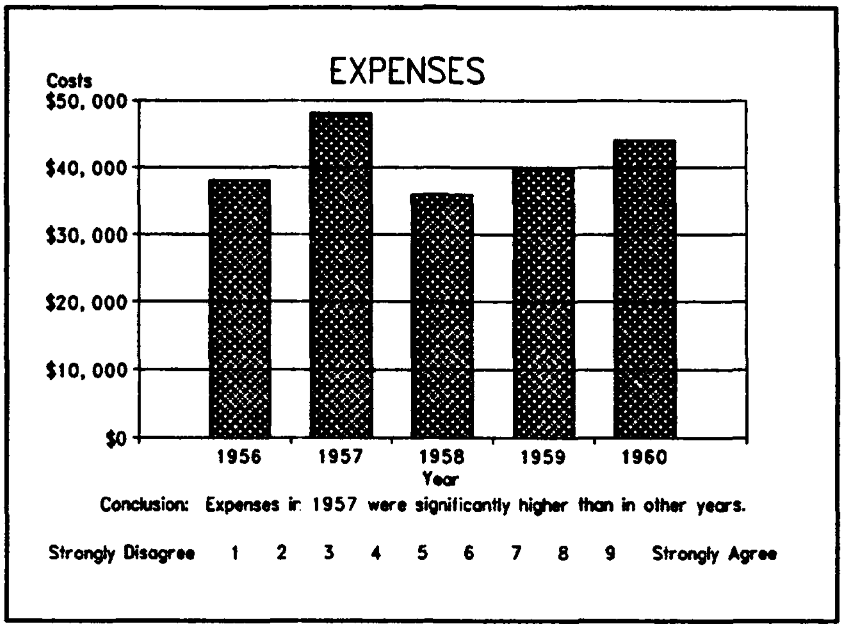
Control Graph 1



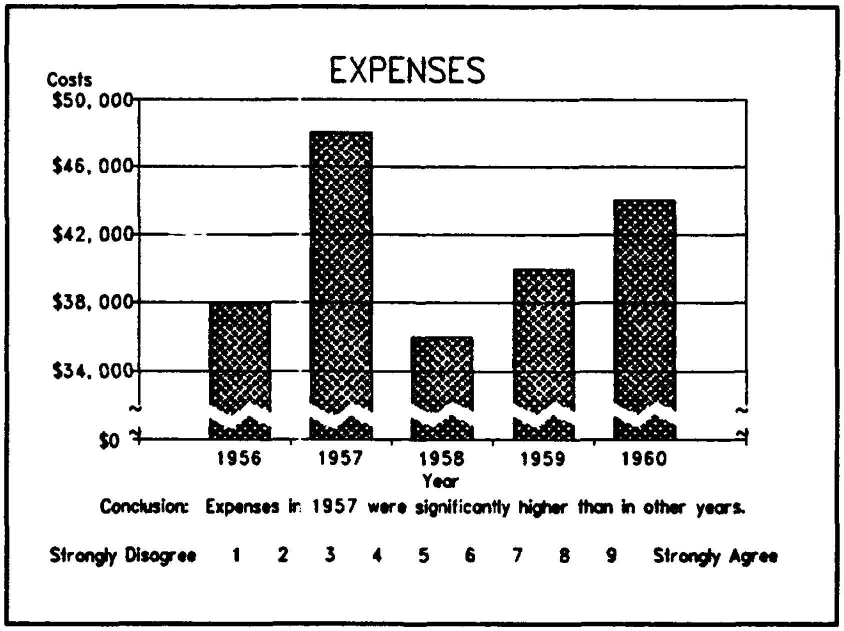
Conclusion: Annual sales increased significantly from 1954 to 1956.

Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree

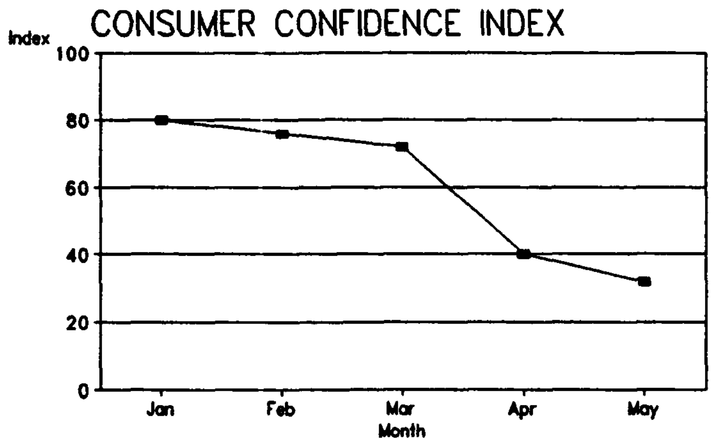
Experimental Graph 1



Control Graph 2



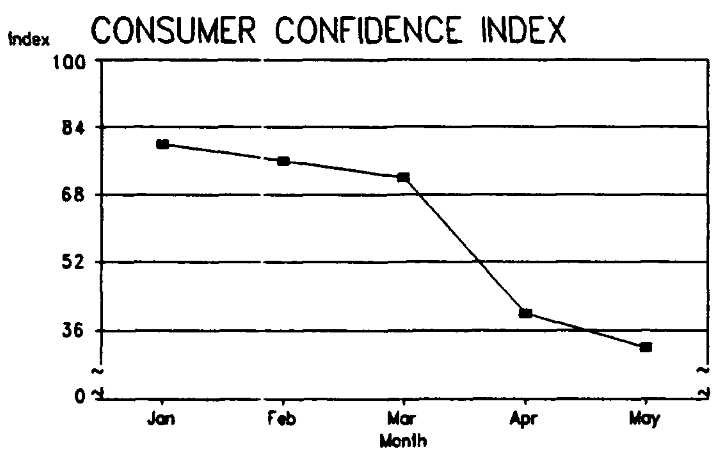
Experimental Graph 2



Conclusion: The consumer confidence index remained relatively constant for the months shown.

Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree

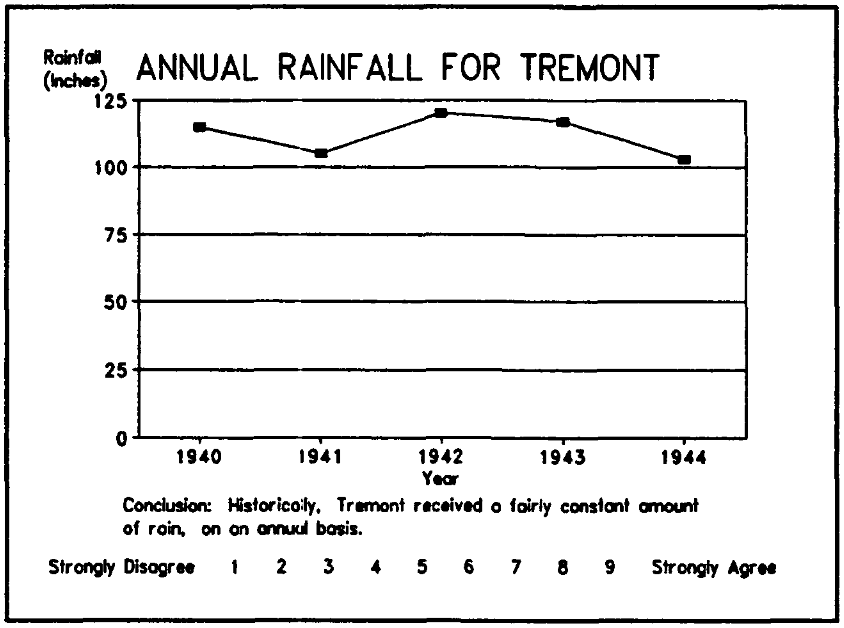
Control Graph 3



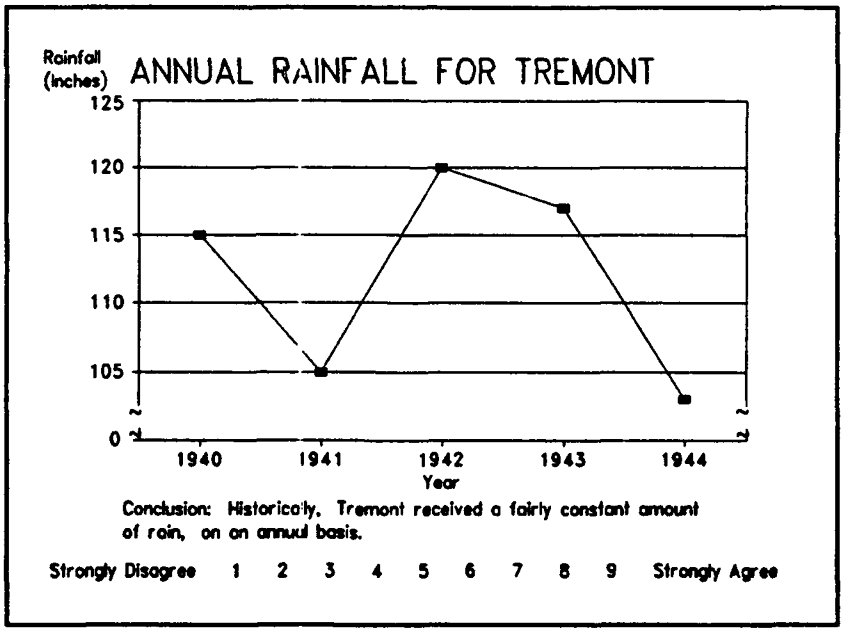
Conclusion: The consumer confidence index remained relatively constant for the months shown.

Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree

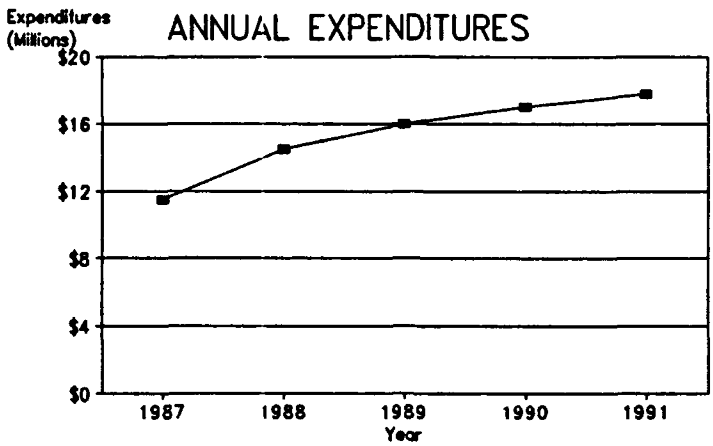
Experimental Graph 3



Control Graph 4



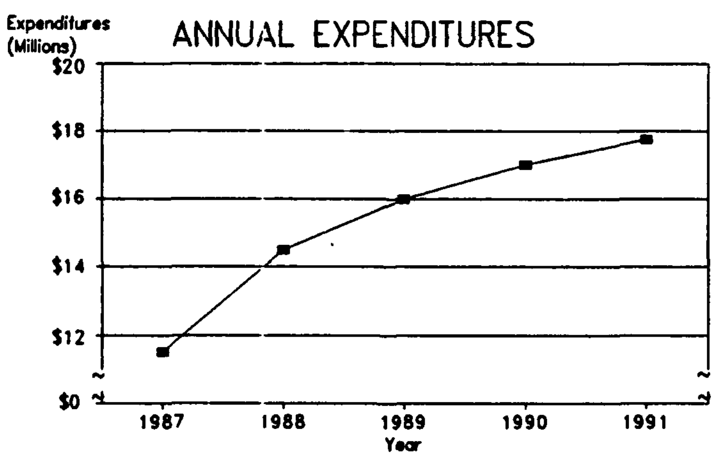
Experimental Graph 4



Conclusion: Annual expenditures are beginning to stabilize.

Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree

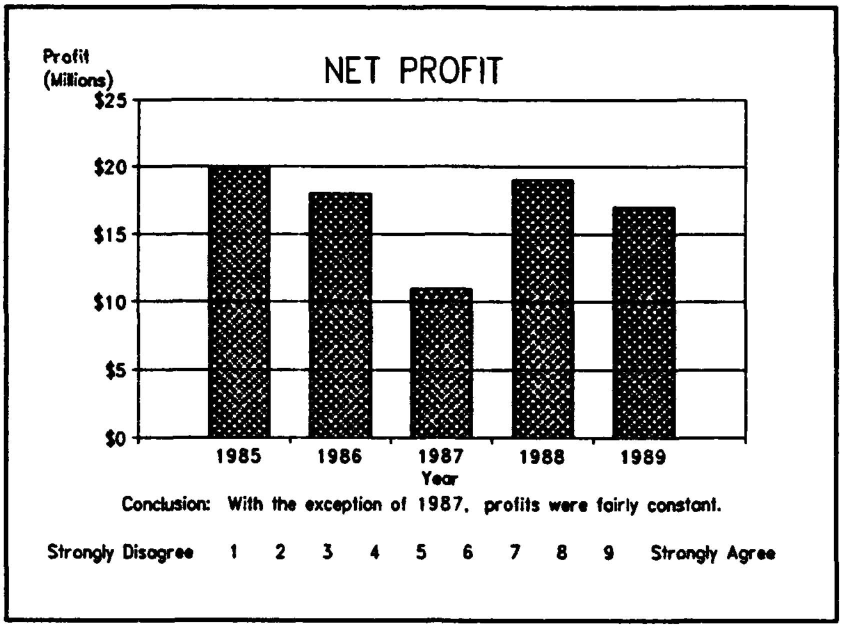
Control Graph 5



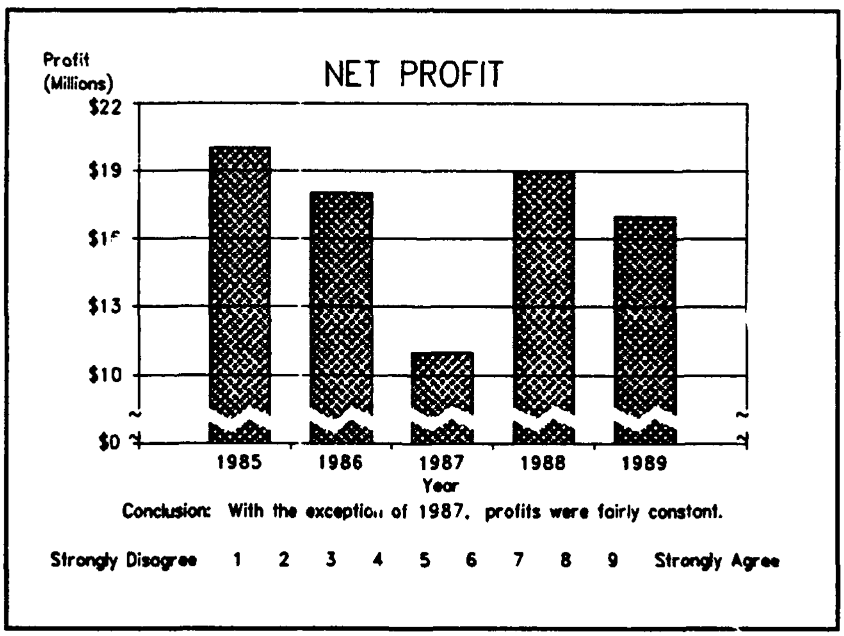
Conclusion: Annual expenditures are beginning to stabilize.

Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree

Experimental Graph 5



Control Graph 6



Experimental Graph 6

DEMOGRAPHIC DATA

1. Sex:

1.____ Male   2.____ Female

2. Age:

1.____ Under 20   2.____ 20 to 29   3.____ 30 to 39

4.____ 40 to 49   5.____ 50 to 59   6.____ Over 59

3. Educational Level (Please specify the highest level obtained):

1.____ High School Diploma   5.____ Bachelors Plus

2.____ High School Plus   6.____ Masters Degree

3.____ Associate Degree   7.____ Masters Plus

4.____ Bachelors Degree   8.____ Doctorate Degree

4. Which of the following fields do you consider to be the primary basis for your professional experience:

1.____ Technical   2.____ Managerial   3.____ Scientific

4.____ Contracts   5.____ Engineering   6.____ Operations

7.____ Support   8.____ Other: ______________________

5. Number of years of Federal Employment:

1.____ 0 to 5   4.____ 16 to 20   7.____ Over 30

2.____ 6 to 10   5.____ 21 to 25   8.____ None

3.____ 11 to 15   6.____ 26 to 30

6. How often do you use graphs in decision-making?

1.____ Every day   2.____ Every other day

3.____ Once a week   4.____ Once a Month

5.____ Once a year   6.____ Never

7. How often do you construct graphs for presentations?

1.____ Every day   2.____ Every other day

3.____ Once a week   4.____ Once a Month

5.____ Once a year   6.____ Never

8. Do you have any comments about the content or presentation of this experiment? If so, please write them on the back of this page.

Appendix C: How to Use Statistix

This appendix describes the specific steps necessary to conduct the statistical tests used in this research in Statistix. Appendix D describes the terms and abbreviations used in the Statistix output tables.

To run Statistix, change to the "SX" subdirectory by typing "cd\sx", then type "SX" to start the program. This will display the base menu.

Statistix is a menu driven program. There are two ways to select items on the menu. The first way is to use the up and down arrow keys until the desired item is highlighted, and then pressing the "return" key. The second way to select a menu item is press the highlighted letter of the menu item desired (each menu item has one character in the name of the item that is highlighted). The keys "esc", "F1", and "return" have special purposes. The esc key will take you back one level in the program. If you are in one of the sub-level menus and you press esc, you will go back to the main menu. The F1 key executes the selected routine. The return key will accept the data entered, or advance the program to the next level.

The first step in conducting any statistical procedure is to enter the data. At the base menu, type "d" for the Data Management sub-menu, then type "d" again for data entry. At this point, the data entry panel appears. This is where you assign an identifier to the data set and define the variable names. The variable names: CGROUP, CP1, CP2, CC1...EE2, EAGE from Tables 25 and 26 were entered (separated by a space) in this panel. Pressing F1 will causes a spreadsheet type screen to appear where the variable values or observations are entered. To enter the data, you simply type the desired value, and then press the down arrow to go to the next case. The data contained in Tables 25 and 26 was entered in this fashion. Once all of the data has been entered press "esc".

Before the statistical tests can be conducted some transformations of the variables must be performed. To do this, from the data management sub-menu press "T" to activate the transformation panel. As described above, the pretest scores and the posttest scores must each be totaled, and the difference obtained. These calculations were done by typing the following information:

CPRETEST=CP1+CP2

CPOSTTEST=CC1+CC2

EPRETEST=EP1+EP2

EPOSTTEST=EE1+EE2

CDELTA=CPRETEST-CPOSTTEST

EDELTA=EPRETEST-EPOSTTEST

CDELTA1=CP1-CC1

EDELTA1=EP1-EE1

After each line is entered "F1" is pressed to perform the transformation. Statistix automatically creates a new set of variable values defined by the above formulas and enters them into the data set. The transformations CDELTA1 and EDELTA1 will be used later in the ANOVA example. The calculations resulting from the transformations are contained in Table 27.

All the initial steps have now been accomplished, and the two-sample t test can be conducted. From the base menu, press "O" to select the One, two & multi-sample tests sub-menu. Next press "T" for the two sample t test. At this point, the two-sample t test panel will appear. The data for this test can be in one of two formats, categorical or tabular. The delta values in the above data set are in tabular format so type "T", and then press "return". The next step is to identify the names of the two variables to be tested. Type "CDELTA EDELTA" then press "F1". Table 28 is the output from this test. The P-value of .0001 indicates that there was a statistical difference between the control and experimental group responses.

TABLE 25

DUMMY DATA FOR CONTROL GROUP

CASE	CGROUP	CP1	CP2	CC1	CC2	CAGE
1	1	2	7	1	8	1
2	1	3	5	1	9	2
3	1	2	4	2	7	3
4	2	1	3	2	8	2
5	2	1	2	1	9	3
6	2	3	4	3	7	1
7	3	2	3	1	7	4
8	3	1	6	3	8	3
9	3	3	3	2	9	3
10	3	1	1	3	9	2

TABLE 26

DUMMY DATA FOR EXPERIMENTAL GROUP

CASE	EGROUP	EP1	EP2	EE1	EE2	EAGE
1	1	9	4	1	7	1
2	1	8	3	2	5	4
3	1	7	2	3	6	3
4	2	9	3	2	8	3
5	2	8	2	2	9	2
6	2	7	1	1	6	3
7	2	6	3	2	9	1
8	3	8	1	1	8	2
9	3	9	3	2	6	4
10	3	8	2	3	7	2

TABLE 27

DUMMY DATA TRANSFORMATIONS

CASE	CPRETEST	CPOSTTEST	EPRETEST	EPOSTTEST	CDELTA	EDELTA	CDELTA1	EDELTA1
1	9	9	13	8	0	5	1	8
2	8	10	11	7	-2	4	2	6
3	6	9	9	9	-3	0	0	4
4	4	10	12	10	-6	2	-1	7
5	3	10	10	11	-7	-1	0	6
6	7	10	8	7	-3	1	0	6
7	5	8	9	11	-3	-2	1	4
8	7	11	9	9	-4	0	-2	7
9	6	11	12	8	-5	4	1	7
10	2	12	10	10	-10	0	-2	5

To conduct the Rank Sum test in Statistix, type "O" from the base menu, and then type "R". This will activate the Rank Sum panel. The data is tabular, so type "T", and then press "return". Again, type in the variable names "CDELTA EDELTA", then press "F1". The results of the Rank Sum test are contained in Table 29. Once again, the P-value of .0001 suggests that there was a statistical difference between the control and experimental group responses. The result of this non-parametric test confirmed the result of the two-sample t test.

In order to conduct an ANOVA to determine the level of significance for the individual graphs, the difference between the pretest and posttest responses for each group must be determined. This difference was already calculated when the transformations were performed. CDELTA1 is the difference between the control group pretest and posttest responses for graph 1, and EDELTA1 is the difference for the experimental group. To conduct the ANOVA, from the base menu type "O", then type "O" again. At this point, the One Way AOV panel is displayed. AOV is another abbreviation for analysis of variance. Type "T" for tabular data, then type "CDELTA1 EDELTA1" and press "F1". The results of the ANOVA are contained in Table 30. The P-value of .0000 indicates that there was a statistical difference between the control and experimental group responses for this particular graph.

To conduct an ANOVA on the demographic data, at the base menu, type "O" then "O" again. When the One Way AOV panel appears, type "C" then press "return". First the name of the dependent variable (the observed values) is requested, so type "CDELTA" then press "return". Next the name of the categorical variable is requested, so type "CAGE" then press "F1". The results of this ANOVA are contained in Table 31. The P-value of .3910 indicates that there was no statistical difference between the responses of members of the control group when the subjects are segregated on the basis of their age.

TABLE 28

TWO SAMPLE T TESTS FOR CDELTA VS EDELTA

VARIABLE	MEAN	SAMPLE SIZE	S.D.	S.E.
CDELTA	-4.300	10	2.830	8.950E-01
EDELTA	1.300	10	2.359	7.461E-01

	T	DF	P
EQUAL VARIANCES	-4.81	18	0.0001
UNEQUAL VARIANCES	-4.81	17.4	0.0002

	F	NUM DF	DEN DF	P
TESTS FOR EQUALITY OF VARIANCES	1.44	9	9	0.2982

CASES INCLUDED 20   MISSING CASES 0

TABLE 29

RANK SUM TWO SAMPLE (MANN-WHITNEY) TEST FOR CDELTA VS EDELTA

VARIABLE	RANK SUM	SAMPLE SIZE	U STAT	AVERAGE RANK
CDELTA	59.00	10	4.000	5.9
EDELTA	151.0	10	96.00	15.1
TOTAL	210.0	20		

EXACT PROBABILITY OF A RESULT AS OR MORE EXTREME
THAN THE OBSERVED RANKS (1 TAILED P VALUE)   0.0001

NORMAL APPROXIMATION WITH CONTINUITY CORRECTION   3.439
TWO TAILED P VALUE FOR NORMAL APPROXIMATION   0.0006

TOTAL NUMBER OF VALUES WHICH WERE TIED   11
MAX. DIFF. ALLOWED BETWEEN TIES   1.0E-0005

CASES INCLUDED 20   MISSING CASES 0

TABLE 30

ONE WAY ANOVA FOR: CDELTA1 EDELTA1

SOURCE	DF	SS	MS	F	P
BETWEEN	1	180.0	180.0	101.25	0.0000
WITHIN	18	32.00	1.778		
TOTAL	19	212.0			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	0.00	1	1.0000

COCHRAN'S Q	0.5000
LARGEST VAR / SMALLEST VAR	1.000

COMPONENT OF VARIANCE FOR BETWEEN GROUPS	17.82
EFFECTIVE CELL SIZE	10.0

VARIABLE	MEAN	SAMPLE SIZE	GROUP STD DEV
CDELTA1	0.000	10	1.333
EDELTA1	6.000	10	1.333
TOTAL	3.000	20	1.333

CASES INCLUDED 20 MISSING CASES 0

TABLE 31

ONE WAY ANOVA FOR CDELTA = CAGE

SOURCE	DF	SS	MS	F	P
BETWEEN	3	26.85	8.950	1.19	0.3910
WITHIN	6	45.25	7.542		
TOTAL	9	72.10			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	1.46	2	0.4826

COCHRAN'S Q	0.6833
LARGEST VAR / SMALLEST VAR	5.486

COMPONENT OF VARIANCE FOR BETWEEN GROUPS	5.951E-01
EFFECTIVE CELL SIZE	2.4

CAGE	MEAN	SAMPLE SIZE	GROUP STD DEV
1	-1.500	2	2.121
2	-6.000	3	4.000
3	-4.750	4	1.708
4	-3.000	1	M
TOTAL	-4.300	10	2.746

CASES INCLUDED 10 MISSING CASES 0

Appendix D: Description of Terms and Variables

This appendix contains two tables that describe the terms and variables used in the Statistix output tables. Table 32 provides a description of the terms, and identifies which type of output table the terms appear in. The descriptions of the terms were compiled from the Statistix User's Manual, and statistics books by Devore and Kachigan (Statistix User's Manual, 1991:87, 183, 215; Devore, 1991:25, 315, 378; Kachigan, 1991:102). The "D S" in the Tables Affected column of Table 52 is an abbreviation for descriptive statistics.

Table 33 contains a description of all of the variable names that are used in this experiment. The first letter in each variable name (C or E) designates the subject's group (control or experimental). In Table 53, when the terms dramatic or nondramatic are used it is in reference to the lie factor value associated with the graph(s) in question (the lie factor greater than eight is dramatic).

TABLE 32

DESCRIPTION OF STATISTIX TERMS

TERM	TERMINOLOGY	DESCRIPTION	TABLES AFFECTED
CHI SQ	Chi Square Test	Goodness of fit test. Useful for analyzing two dimensional tables of discrete data.	ANOVA
COCHRAN'S Q	-	Ratio of the largest within group variance over the sum of all within group variances.	ANOVA
DF	number of degrees of freedom	The number of observations in the data collection that are free to vary after the sample statistics have been calculated. Parameter of the T and F statistics.	ANOVA, T test
F	F statistic	The test statistic for ANOVA (see Chapter III.).	ANOVA, T test
DEN DF	denominator degrees of freedom	The degrees of freedom in the denominator of the F statistic (MSE).	T test
MS	mean square	The two F statistic sum of squares divided by their appropriate degrees of freedom.	ANOVA
NUM DF	numerator degrees of freedom	The degrees of freedom in the numerator of the F statistic (MSTr).	T test
P	P-value	The smallest level of significance at which Ho would be rejected when a specified test procedure is used on a given data set.	ANOVA, T test, Rank Sum
RANK SUM	-	Non-parametric test of differences in the central values of samples from two independent samples.	Rank Sum
SS	sum of squares	There are three sum of squares, SST is a measure of the total deviation in the data, SSE is the measure of variation present even if Ho is true, SSTr is the amount of variation due to differences in the average values. The sum of squares values are used to calculate the mean square.	ANOVA

TABLE 32 CONTINUED

TERM	**TERMINOLOGY**	**DESCRIPTION**	**TABLES AFFECTED**
STD DEV or S.D.	standard deviation	A measure of the amount of variability or dispersion about the mean, present in a data set. The square root of the variance.	ANOVA, T test, D S *
S.E.	standard error	Tells roughly within what distance of an estimator the true value can be expected to lie.	ANOVA, D S *, T test,
T	T statistic	The test statistic of the t test (see Chapter III.).	T test
U STAT	U statistic	The Mann-Whitney test statistic which is mathematically equivalent to the Rank Sum test statistic w.	Rank Sum

* D S stands for Descriptive Statistics

TABLE 33

DESCRIPTION OF EXPERIMENTAL VARIABLES

VARIABLE	DESCRIPTION
CAGE (EGAE)	Age of subject. Demographic factor.
CBT (EBT)	Total difference (gain score) between the posttest and associated pretest bar graphs.
CBTM (EBTM)	Total difference (gain score) between the posttest and associated pretest bar graphs, with the most dramatic bar graph excluded.
CCONST (ECONST)	How often the subject constructs graphs. Demographic factor.
CC1 (EE1)	Subject's response to posttest graph 1.
CC2 (EE2)	Subject's response to posttest graph 2.
CC3 (EE3)	Subject's response to posttest graph 3.
CC4 (EE4)	Subject's response to posttest graph 4.
CC5 (EE5)	Subject's response to posttest graph 5.
CC6 (EE6)	Subject's response to posttest graph 6.
CDELTA (EDELTA)	Total difference (gain score) between the total of the pretest responses and the total of the posttest responses.
CEDLVL (EEDLVL)	The education level achieved by the subject. Demographic factor.
CEMP (EEMP)	Subject's years of Federal Service. Demographic factor.
CGROUP (EGROUP)	The subject's administrative group number. Demographic factor.
CLT (ELT)	Total difference (gain score) between the posttest and associated pretest line graphs.
CLTM (ELTM)	Total difference (gain score) between the posttest and associated pretest line graphs, with the most dramatic line graph excluded.
CNSIGN (ESIGN)	The total of the gain scores of the nondramatic graphs, both bar and line.
CP1 (EP1)	Subject's response to pretest graph 1.
CP2 (EP2)	Subject's response to pretest graph 2.
CP3 (EP3)	Subject's response to pretest graph 3.
CP4 (EP4)	Subject's response to pretest graph 4.
CP5 (EP5)	Subject's response to pretest graph 5.

TABLE 33 CONTINUED

VARIABLE	DESCRIPTION
CP6 (EP6)	Subject's response to pretest graph 6.
CSEX (ESEX)	The subject's sex. Demographic factor.
CSIGN (ESIGN)	The total of the gain scores of the dramatic graphs, both bar and line.
CUSE (EUSE)	How often the subject uses graphs in decision making. Demographic factor.
C1T (E1T)	The difference between posttest response 1 and its associated pretest response (see Table 21).
C2T (E2T)	The difference between posttest response 2 and its associated pretest response (see Table 21).
C3T (E3T)	The difference between posttest response 3 and its associated pretest response (see Table 21).
C4T (E4T)	The difference between posttest response 4 and its associated pretest response (see Table 21).
C5T (E5T)	The difference between posttest response 5 and its associated pretest response (see Table 21).
C6T (E6T)	The difference between posttest response 6 and its associated pretest response (see Table 21).

Appendix E: Experimental Results

This appendix contains all of the results of the experiment. The control group results are contained in Table 34, the subject's responses to the graphs are contained in Table 34A, and the subject's demographic data is contained in Table 34B. The experimental group results are contained in Table 35 (35A is the graph responses and 35B is the demographic data).

TABLE 34A

CONTROL GROUP DATA

CASE	GROUP	P1	P2	P3	P4	P5	P6	PRETEST TOTAL	C1	C2	C3	C4	C5	C6	POSTTEST TOTAL	DELTA DIFF
1	1	3	8	6	7	2	6	32	5	6	2	6	6	7	32	0
2	1	3	8	1	4	7	3	26	7	4	1	8	5	8	33	7
3	1	1	9	1	2	4	3	20	8	9	1	7	7	8	40	20
4	1	1	8	8	2	4	7	30	7	6	2	8	6	7	36	6
5	1	4	8	6	3	4	5	30	7	8	3	7	6	6	37	7
6	1	5	9	1	3	5	6	29	7	7	1	6	1	9	31	2
7	1	1	8	5	2	3	8	27	7	4	1	8	2	3	25	-2
8	1	3	8	8	7	3	7	36	8	8	1	3	8	7	35	-1
9	1	2	9	8	1	4	3	27	8	4	1	8	7	7	35	8
10	1	2	7	3	2	6	3	23	7	3	3	7	7	7	34	11
11	1	6	6	8	7	7	6	40	8	7	3	8	3	8	37	-3
12	1	3	8	6	3	8	7	35	8	8	2	8	3	4	33	-2
13	1	3	6	3	3	4	4	23	5	3	2	8	8	7	33	10
14	1	2	9	8	4	4	7	34	8	8	1	6	5	6	34	0
15	1	2	7	7	2	2	1	21	1	7	3	3	5	3	22	1
16	1	1	9	8	8	3	7	36	2	6	1	1	4	8	22	-14
17	2	2	8	7	1	5	7	30	2	7	1	7	7	4	28	-2
18	2	8	9	1	9	8	3	38	7	9	1	6	6	7	36	-2
19	2	1	8	7	2	4	3	25	7	5	2	6	7	7	34	9
20	2	5	9	9	9	7	4	43	5	6	1	8	1	8	29	-14
21	2	5	9	8	1	7	3	33	4	6	1	8	5	6	30	-3
22	2	2	9	3	8	3	5	30	3	6	1	8	7	5	30	0
23	2	3	5	6	3	6	2	25	6	3	2	4	3	3	21	-4
24	2	1	9	1	8	5	5	29	8	9	1	8	8	3	37	8
25	2	3	8	9	3	6	3	32	6	7	2	4	8	8	35	3
26	2	2	8	8	2	2	3	25	7	4	2	7	7	5	32	7
27	2	1	8	8	2	7	4	30	6	5	1	8	5	6	31	1
28	2	3	9	9	1	4	3	29	9	9	1	7	7	5	38	9
29	2	2	9	7	1	9	7	35	8	9	1	9	6	9	42	7
30	2	3	8	1	3	3	3	21	7	7	1	9	6	8	38	17
31	2	2	8	8	2	6	2	28	6	7	1	7	2	7	30	2
32	2	3	8	8	2	5	3	29	8	9	2	8	4	8	39	10
33	2	1	8	7	2	4	2	24	6	6	1	3	3	7	26	2
34	2	2	9	9	1	7	4	32	5	6	1	[illegible]	7	8	35	3
35	2	3	9	1	2	8	6	29	1	7	2	9	2	9	30	1
36	2	3	7	6	2	8	3	29	6	6	3	7	8	8	38	9
37	2	7	8	7	4	7	6	39	7	7	3	8	3	8	36	-3
38	2	3	9	7	3	5	6	33	9	9	2	7	5	8	40	7
39	3	3	9	2	1	4	4	23	2	7	1	7	7	4	28	5
40	3	2	8	9	1	7	1	28	5	9	1	8	8	7	38	10
41	3	2	8	6	7	8	2	33	3	7	1	8	4	6	29	-4
42	3	1	8	9	3	3	3	27	2	9	1	6	9	8	35	8
43	3	4	8	1	1	7	3	24	7	6	2	8	2	8	33	9
44	3	3	8	3	4	4	3	25	5	3	1	7	6	7	29	4
45	3	1	8	1	1	3	1	15	6	8	1	4	1	3	23	8
46	3	3	5	2	9	7	6	32	1	6	5	7	2	5	26	-6

TABLE 34A CONTINUED

CASE	GROUP	P1	P2	P3	P4	P5	P6	PRETEST TOTAL	C1	C2	C3	C4	C5	C6	POSTTEST TOTAL	DELTA DIFF
47	3	1	9	1	1	5	1	18	9	7	1	3	1	9	30	12
48	3	1	9	8	1	7	8	34	9	7	2	8	7	8	41	7
49	3	4	9	4	1	6	3	27	4	4	1	7	3	8	27	0
50	3	7	9	1	2	9	4	32	6	6	2	8	4	9	35	3
51	3	2	8	1	3	5	3	22	6	7	2	8	6	7	36	14
52	3	3	9	1	1	6	2	22	6	7	1	6	8	6	34	12
53	3	3	8	4	3	4	3	25	6	4	1	7	4	6	28	3
54	3	1	7	1	6	7	3	25	8	6	1	7	4	7	33	8
55	3	4	8	1	2	7	2	24	6	6	2	8	3	7	32	8
56	3	9	8	9	2	7	3	38	8	9	3	7	3	8	38	0
57	4	1	9	9	3	9	1	32	8	9	1	5	1	1	25	-7
58	4	2	7	4	8	7	3	31	6	7	1	8	7	8	37	6
59	4	2	8	8	2	4	4	28	3	7	1	8	3	7	29	1
60	4	1	9	9	1	1	7	28	9	7	1	7	1	3	28	0
61	4	2	9	9	1	7	1	29	8	6	1	8	8	7	38	9
62	4	5	9	8	1	6	5	34	6	6	1	9	5	8	35	1
63	4	2	9	9	1	7	3	31	7	8	1	4	7	7	34	3
64	4	3	9	9	2	8	2	33	3	9	1	7	9	8	37	4
65	4	2	8	1	1	8	3	23	7	7	3	8	5	8	38	15
66	4	7	8	7	8	6	7	43	8	5	1	8	8	8	38	-5
67	4	2	8	1	3	7	3	24	7	7	2	7	5	6	34	10
68	4	6	9	1	7	7	6	36	8	7	1	5	8	7	36	0
69	4	7	9	1	6	8	7	38	7	6	1	5	5	8	32	-6
70	4	3	8	9	5	7	7	39	9	8	2	9	8	7	43	4
71	4	1	8	1	1	7	3	21	7	8	1	8	8	8	40	19
72	4	4	9	6	9	5	8	41	7	6	2	3	4	8	30	-11
73	4	2	8	1	5	7	3	26	7	4	1	5	5	7	29	3
74	4	3	9	5	1	9	7	34	6	7	1	8	4	7	33	-1

TABLE 34B

CONTROL GROUP DEMOGRAPHIC DATA

CASE	GROUP	TIME	SEX	AGE	EDUCATION LEVEL	PROFESSIONAL EXPERIENCE	FEDERAL EMPLOYMENT	GRAPH USE	GRAPH CONSTRUCTION
1	1	15	1	2	5	8	2	4	4
2	1	15	1	2	5	2	2	4	4
3	1	15	1	3	5	5	3	2	4
4	1	15	1	2	5	2	2	3	3
5	1	15	1	3	5	2	1	4	4
6	1	15	1	3	5	2	2	4	4
7	1	15	2	3	5	5	3	4	4
8	1	15	1	3	4	2	2	4	4
9	1	15	1	3	5	8	3	3	4
10	1	15	1	2	5	2	2	3	3
11	1	15	1	3	5	8	3	4	4
12	1	15	1	3	5	8	3	3	3
13	1	15	1	3	5	1	3	4	4
14	1	15	1	2	5	8	2	3	3
15	1	15	1	3	5	1	2	4	4
16	1	15	1	2	5	2	2	4	4
17	2	15	1	2	4	5	1	4	4
18	2	15	1	2	6	1	1	4	4
19	2	15	1	2	6	7	1	2	2
20	2	15	1	4	6	6	4	3	4
21	2	15	2	2	6	7	1	4	4
22	2	15	1	2	4	5	1	2	3
23	2	15	1	3	5	8	2	4	3
24	2	15	1	3	4	1	3	5	4
25	2	15	2	5	6	1	6	6	5
26	2	15	1	2	4	7	1	1	1
27	2	15	1	3	6	2	2	5	5
28	2	15	2	3	2	6	3	6	6
29	2	15	1	3	5	1	2	4	4
30	2	15	1	2	6	7	1	3	3
31	2	15	2	2	4	7	1	4	4
32	2	15	2	3	7	8	2	5	5
33	2	15	1	2	4	7	1	3	3
34	2	15	2	3	1	8	3	6	6
35	2	15	2	2	4	8	2	6	6
36	2	15	1	4	6	6	5	3	4
37	2	15	2	3	4	1	3	4	4
38	2	15	2	3	3	8	5	3	3
39	3	15	1	2	4	2	1	4	4
40	3	15	1	2	4	2	1	4	4
41	3	15	1	2	4	2	1	3	5
42	3	15	2	2	4	2	1	6	4
43	3	15	2	2	4	2	1	5	5
44	3	15	2	2	4	4	1	4	4
45	3	15	2	2	4	2	1	4	4
46	3	15	2	2	4	8	1	4	4

TABLE 34B CONTINUED

CASE	GROUP	TIME	SEX	AGE	EDUCATION LEVEL	PROFESSIONAL EXPERIENCE	FEDERAL EMPLOYMENT	GRAPH USE	GRAPH CONSTRUCTION
47	3	15	1	2	4	1	1	3	6
48	3	15	2	2	4	2	1	5	4
49	3	15	1	2	4	8	1	6	6
50	3	15	2	2	4	2	1	6	4
51	3	15	2	2	4	2	8	5	4
52	3	15	1	2	4	8	1	3	3
53	3	15	1	2	4	2	1	5	5
54	3	15	2	2	4	2	1	3	4
55	3	15	1	2	4	2	1	4	4
56	3	15	2	2	4	2	1	5	5
57	4	30	2	2	4	2	1	3	4
58	4	30	1	2	4	1	8	4	4
59	4	30	1	2	4	2	8	4	4
60	4	30	2	2	4	2	8	3	4
61	4	30	2	2	4	8	1	4	4
62	4	30	2	2	4	8	1	6	4
63	4	30	2	2	4	1	8	6	6
64	4	30	1	2	4	2	1	4	2
65	4	30	2	2	4	2	8	5	4
66	4	30	1	2	4	8	8	4	4
67	4	30	1	2	4	8	1	3	3
68	4	30	1	2	4	8	8	4	4
69	4	30	1	2	4	8	1	4	4
70	4	30	1	2	4	2	2	6	5
71	4	30	1	2	4	8	8	2	4
72	4	30	1	2	4	8	1	5	4
73	4	30	2	2	4	1	1	5	5
74	4	30	2	2	4	2	1	4	4

TABLE 35A

EXPERIMENTAL GROUP DATA

CASE	GROUP	P1	P2	P3	P4	P5	P6	PRETEST TOTAL	E1	E2	E3	E4	E5	E6	POSTTEST TOTAL	DELTA DIFF
1	1	1	1	9	1	9	5	26	8	6	1	6	6	5	32	6
2	1	2	9	7	7	3	7	35	8	8	1	2	3	3	25	-10
3	1	2	9	8	3	8	3	33	8	3	2	9	8	3	33	0
4	1	4	7	2	3	7	7	30	5	4	3	7	6	7	32	2
5	1	2	9	1	2	8	4	26	8	3	2	7	3	7	30	4
6	1	1	9	8	7	7	4	36	7	9	1	3	7	8	35	-1
7	1	4	9	8	1	6	2	30	8	7	2	5	6	7	35	5
8	1	3	8	8	8	8	3	38	7	8	2	2	2	3	24	-14
9	1	1	6	5	6	6	4	28	8	9	3	3	7	7	37	9
10	1	3	7	7	5	3	4	29	7	8	2	3	2	4	26	-3
11	1	2	9	8	5	7	6	37	7	9	2	8	7	8	41	4
12	1	3	9	1	4	4	3	24	9	5	3	3	5	8	33	9
13	1	4	9	2	1	8	4	28	9	7	1	3	8	4	32	4
14	1	1	9	4	7	8	9	38	9	7	1	3	3	4	27	-11
15	1	6	8	7	2	7	6	36	3	3	3	5	7	7	28	-8
16	1	1	9	1	8	1	2	22	9	9	1	1	8	8	36	14
17	2	3	7	6	8	1	1	26	3	7	1	1	7	5	24	-2
18	2	4	9	7	6	8	7	41	3	1	1	2	2	2	11	-30
19	2	2	9	9	9	5	7	41	1	4	1	7	1	9	23	-18
20	2	2	7	1	3	3	3	19	1	7	5	1	6	3	23	4
21	2	3	9	9	2	7	3	33	8	7	1	5	5	5	31	-2
22	2	3	8	7	4	4	5	31	4	7	4	5	2	4	26	-5
23	2	3	9	7	2	7	3	31	9	9	1	1	8	8	36	5
24	2	3	7	6	2	6	4	28	7	7	2	4	4	5	29	1
25	2	1	9	8	1	9	5	33	9	7	1	9	9	8	43	10
26	2	3	8	8	2	8	7	36	6	8	2	7	4	7	34	-2
27	2	6	8	8	2	4	3	31	8	9	2	3	5	3	30	-1
28	2	3	8	1	3	6	6	27	8	9	1	4	6	8	36	9
29	2	1	8	9	3	3	2	26	9	8	1	1	4	6	29	3
30	2	3	4	9	6	1	5	28	7	7	1	3	4	7	29	1
31	2	1	8	6	2	5	2	24	3	8	2	3	2	4	22	-2
32	2	1	9	9	1	9	7	36	8	7	1	3	1	8	28	-8
33	2	3	8	9	4	4	6	34	7	9	2	4	5	3	30	-4
34	2	4	8	2	8	7	3	32	6	6	1	2	7	8	30	-2
35	2	6	8	9	5	6	6	40	7	8	3	7	5	7	37	-3
36	2	2	9	2	2	6	6	27	7	7	1	4	3	7	29	2
37	2	5	9	9	1	8	5	37	9	9	1	7	7	9	42	5
38	3	2	9	8	5	4	3	31	8	9	2	1	2	4	26	-5
39	3	4	9	1	4	8	7	33	9	9	1	6	3	8	36	3
40	3	2	8	1	3	4	5	23	7	9	1	2	6	6	31	8
41	3	1	9	9	2	4	1	26	7	8	1	6	8	9	39	13
42	3	2	8	9	1	3	3	26	9	4	1	4	8	9	35	9
43	3	4	8	9	6	7	4	38	7	8	1	5	7	3	31	-7
44	3	1	8	9	1	9	8	36	9	6	1	1	1	9	27	-9
45	3	6	9	2	1	4	2	24	1	9	3	5	7	7	32	8
46	3	3	7	9	1	5	6	31	7	9	1	4	5	6	32	1

TABLE 35A CONTINUED

CASE	GROUP	P1	P2	P3	P4	P5	P6	PRETEST TOTAL	E1	E2	E3	E4	E5	E6	POSTTEST TOTAL	DELTA DIFF
47	3	1	9	6	1	2	1	20	8	9	1	1	7	4	30	10
48	3	2	9	8	4	5	3	31	9	3	1	1	8	6	28	-3
49	3	2	6	8	3	7	3	29	8	8	4	3	4	8	35	6
50	3	3	8	9	1	8	6	35	9	9	1	2	3	8	32	-3
51	3	9	9	1	1	9	2	31	6	9	1	1	7	6	30	-1
52	3	3	7	9	5	4	4	32	7	9	3	3	6	7	35	3
53	3	1	8	9	3	6	4	31	8	5	1	4	4	7	29	-2
54	3	1	9	2	1	9	9	31	9	9	2	1	9	9	39	8
55	3	4	9	7	3	4	6	33	2	9	1	2	7	7	28	-5
56	3	2	7	3	2	7	5	26	8	8	2	3	4	7	32	6
57	4	3	8	2	3	6	3	25	8	4	1	6	7	6	32	7
58	4	2	7	1	5	6	3	24	8	8	1	1	6	4	28	4
59	4	3	8	9	2	8	6	36	7	8	1	2	7	7	32	-4
60	4	1	8	7	1	3	5	25	8	7	1	1	2	1	20	-5
61	4	7	9	8	1	8	8	41	2	6	1	4	3	9	25	-16
62	4	2	9	1	1	7	2	22	9	7	1	1	8	7	33	11
63	4	1	8	6	3	8	3	29	8	7	2	7	7	4	35	6
64	4	6	9	3	3	7	3	31	7	7	2	1	7	8	32	1
65	4	4	9	4	1	7	4	29	4	9	2	2	3	7	27	-2
66	4	4	9	2	1	4	3	23	5	7	3	6	7	8	36	13
67	4	3	8	7	6	5	4	33	4	6	3	6	7	7	33	0
68	4	1	3	9	2	8	5	28	9	2	1	7	7	4	30	2
69	4	3	8	2	1	4	1	19	4	5	1	3	6	3	22	3
70	4	2	7	2	7	3	2	23	8	8	2	6	6	4	34	11
71	4	1	8	9	4	6	1	29	7	3	1	2	8	4	25	-4
72	4	1	8	9	1	5	2	26	9	4	1	1	1	6	22	-4
73	4	2	9	2	1	8	4	26	8	8	2	6	7	8	39	13

TABLE 35B

EXPERIMENTAL GROUP DEMOGRAPHIC DATA

CASE	GROUP	TIME	SEX	AGE	EDUCATION LEVEL	PROFESSIONAL EXPERIENCE	FEDERAL EMPLOYMENT	GRAPH USE	GRAPH CONSTRUCTION
1	1	15	1	3	4	2	2	6	3
2	1	15	1	3	5	5	3	4	4
3	1	15	1	3	5	5	3	1	4
4	1	15	1	2	5	8	2	1	2
5	1	15	1	3	5	8	2	4	4
6	1	15	1	2	5	2	3	1	4
7	1	15	1	3	5	2	2	3	4
8	1	15	1	3	5	2	3	2	3
9	1	15	1	2	5	2	2	1	3
10	1	15	1	3	4	5	4	6	6
11	1	15	1	2	5	2	2	1	3
12	1	15	1	3	5	8	3	1	3
13	1	15	1	3	5	8	2	4	4
14	1	15	1	3	7	1	3	2	3
15	1	15	1	3	5	5	2	5	5
16	1	15	1	2	5	5	2	2	3
17	2	15	2	4	3	8	5	4	4
18	2	15	2	3	4	4	2	2	3
19	2	15	2	3	5	7	3	4	4
20	2	15	2	3	5	8	8	6	6
21	2	15	1	2	4	8	1	3	3
22	2	15	1	3	6	8	3	3	4
23	2	15	2	4	6	8	6	6	6
24	2	15	1	2	4	7	1	3	4
25	2	15	2	4	2	1	4	5	5
26	2	15	1	6	6	5	4	4	4
27	2	15	1	3	6	7	3	4	4
28	2	15	1	3	6	8	3	4	3
29	2	15	1	4	6	8	4	1	1
30	2	15	2	4	3	2	4	3	4
31	2	15	2	4	4	8	2	4	6
32	2	15	2	2	4	2	1	3	2
33	2	15	2	5	6	7	2	4	4
34	2	15	1	4	2	8	5	3	3
35	2	15	1	3	6	6	4	4	5
36	2	15	1	3	6	2	4	6	5
37	2	15	2	4	4	1	2	4	4
38	3	15	2	2	4	2	1	5	4
39	3	15	1	2	4	2	1	4	4
40	3	15	1	2	4	2	1	3	4
41	3	15	1	2	4	2	1	4	6
42	3	15	2	2	4	2	1	3	4
43	3	15	2	2	4	2	1	4	5
44	3	15	1	2	4	4	8	4	4
45	3	15	1	2	4	8	8	2	4
46	3	15	1	2	4	2	1	4	4

TABLE 35B CONTINUED

CASE	GROUP	TIME	SEX	AGE	EDUCATION LEVEL	PROFESSIONAL EXPERIENCE	FEDERAL EMPLOYMENT	GRAPH USE	GRAPH CONSTRUCTION
47	3	15	1	2	4	8	8	1	4
48	3	15	2	2	4	5	1	3	3
49	3	15	2	2	4	2	1	6	4
50	3	15	1	2	4	2	1	4	4
51	3	15	2	2	4	8	1	4	4
52	3	15	2	2	4	2	1	6	4
53	3	15	2	2	4	6	1	4	4
54	3	15	1	2	4	2	1	5	4
55	3	15	1	2	4	8	2	4	4
56	3	15	1	2	4	2	1	6	4
57	4	30	1	2	4	2	1	3	4
58	4	30	1	2	4	2	8	2	3
59	4	30	2	2	4	2	1	4	4
60	4	30	2	2	4	2	1	3	3
61	4	30	1	2	4	4	1	4	4
62	4	30	1	2	4	2	1	3	4
63	4	30	1	2	4	1	1	4	4
64	4	30	2	2	4	2	1	3	4
65	4	30	2	2	4	1	8	4	3
66	4	30	1	2	4	2	1	5	5
67	4	30	1	2	4	2	1	4	4
68	4	30	2	2	4	8	1	6	6
69	4	30	1	2	4	2	1	2	3
70	4	30	1	2	4	2	1	5	4
71	4	30	1	2	4	2	1	4	4
72	4	30	2	2	4	2	1	3	4
73	4	30	1	2	4	1	1	1	3

Appendix F: ANOVA Results

This appendix contains the results of the ANOVAs conducted on the data. Tables 36 to 47 contain the results of the analysis conducted on the responses to the graphs. Tables 48 to 61 contain the results of the analysis conducted on the demographic data.

TABLE 36

ONE WAY ANOVA FOR: C1T E1T

SOURCE	DF	SS	MS	F	P
BETWEEN	1	28.22	28.22	4.78	0.0303
WITHIN	145	855.3	5.899		
TOTAL	146	883.5			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	3.93	1	0.0475

COCHRAN'S Q 0.6152
LARGEST VAR / SMALLEST VAR 1.598
COMPONENT OF VARIANCE FOR BETWEEN GROUPS 3.037E-01
EFFECTIVE CELL SIZE 73.5

VARIABLE	MEAN	SAMPLE SIZE	GROUP STD DEV
C1T	2.027	74	2.132
E1T	1.151	73	2.696
TOTAL	1.592	147	2.429

TABLE 37

ONE WAY ANOVA FOR: C2T E2T

SOURCE	DF	SS	MS	F	P
BETWEEN	1	9.817	9.817	1.29	0.2585
WITHIN	145	1.106E+03	7.629		
TOTAL	146	1.116E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	0.46	1	0.4962

COCHRAN'S Q 0.5400
LARGEST VAR / SMALLEST VAR 1.174
COMPONENT OF VARIANCE FOR BETWEEN GROUPS 2.977E-02
EFFECTIVE CELL SIZE 73.5

VARIABLE	MEAN	SAMPLE SIZE	GROUP STD DEV
C2T	-7.297E-01	74	2.650
E2T	-1.247	73	2.871
TOTAL	-0.986	147	2.762

TABLE 38

ONE WAY ANOVA FOR: C3T E3T

SOURCE	DF	SS	MS	F	P
BETWEEN	1	22.20	22.20	2.19	0.1414
WITHIN	145	1.472E+03	10.15		
TOTAL	146	1.495E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	0.20	1	0.6538

COCHRAN'S Q 0.5264
LARGEST VAR / SMALLEST VAR 1.112
COMPONENT OF VARIANCE FOR BETWEEN GROUPS 1.639E-01
EFFECTIVE CELL SIZE 73.5

VARIABLE	MEAN	SAMPLE SIZE	GROUP STD DEV
C3T	-3.257	74	3.269
E3T	-2.479	73	3.101
TOTAL	-2.871	147	3.187

TABLE 39

ONE WAY ANOVA FOR: C4T E4T

SOURCE	DF	SS	MS	F	P
BETWEEN	1	341.4	341.4	30.89	0.0000
WITHIN	145	1.603E+03	11.05		
TOTAL	146	1.944E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	0.01	1	0.9183

COCHRAN'S Q 0.5060
LARGEST VAR / SMALLEST VAR 1.024
COMPONENT OF VARIANCE FOR BETWEEN GROUPS 4.495
EFFECTIVE CELL SIZE 73.5

VARIABLE	MEAN	SAMPLE SIZE	GROUP STD DEV
C4T	-3.500	74	3.344
E4T	-4.521E-01	73	3.304
TOTAL	-1.986	147	3.324

TABLE 40

ONE WAY ANOVA FOR: C5T E5T

SOURCE	DF	SS	MS	F	P
BETWEEN	1	3.865E-01	3.865E-01	0.04	0.8481
WITHIN	145	1.522E+03	10.50		
TOTAL	146	1.523E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	1.11	1	0.2917

COCHRAN'S Q 0.5619
LARGEST VAR / SMALLEST VAR 1.283
COMPONENT OF VARIANCE FOR BETWEEN GROUPS -1.376E-01
EFFECTIVE CELL SIZE 73.5

VARIABLE	MEAN	SAMPLE SIZE	GROUP STD DEV
C5T	1.514	74	3.433
E5T	1.411	73	3.031
TOTAL	1.463	147	3.240

TABLE 41

ONE WAY ANOVA FOR: C6T E6T

SOURCE	DF	SS	MS	F	P
BETWEEN	1	22.68	22.68	2.80	0.0965
WITHIN	145	1.175E+03	8.107		
TOTAL	146	1.198E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	0.20	1	0.6535

COCHRAN'S Q 0.5264
LARGEST VAR / SMALLEST VAR 1.112
COMPONENT OF VARIANCE FOR BETWEEN GROUPS 1.983E-01
EFFECTIVE CELL SIZE 73.5

VARIABLE	MEAN	SAMPLE SIZE	GROUP STD DEV
C6T	3.514E-01	74	2.921
E6T	1.137	73	2.770
TOTAL	7.415E-01	147	2.847

TABLE 42

ONE WAY ANOVA FOR: CNSIGN ENSIGN

SOURCE	DF	SS	MS	F	P
BETWEEN	1	40.21	40.21	1.42	0.2349
WITHIN	145	4.099E+03	28.27		
TOTAL	146	4.139E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	0.56	1	0.4544

COCHRAN'S Q 0.5440
LARGEST VAR / SMALLEST VAR 1.193
COMPONENT OF VARIANCE FOR BETWEEN GROUPS 1.625E-01
EFFECTIVE CELL SIZE 73.5

VARIABLE	MEAN	SAMPLE SIZE	GROUP STD DEV
CNSIGN	3.635	74	5.079
ENSIGN	2.589	73	5.547
TOTAL	3.116	147	5.317

TABLE 43

ONE WAY ANOVA FOR: CSIGN ESIGN

SOURCE	DF	SS	MS	F	P
BETWEEN	1	195.6	195.6	8.47	0.0042
WITHIN	145	3.348E+03	23.09		
TOTAL	146	3.544E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	0.89	1	0.3459

COCHRAN'S Q 0.5554
LARGEST VAR / SMALLEST VAR 1.249
COMPONENT OF VARIANCE FOR BETWEEN GROUPS 2.348
EFFECTIVE CELL SIZE 73.5

VARIABLE	MEAN	SAMPLE SIZE	GROUP STD DEV
CSIGN	-2.432	74	4.533
ESIGN	-4.740	73	5.066
TOTAL	-3.578	147	4.805

TABLE 44

ONE WAY ANOVA FOR: CLT ELT

SOURCE	DF	SS	MS	F	P
BETWEEN	1	509.3	509.3	25.15	0.0000
WITHIN	145	2.936E+03	20.25		
TOTAL	146	3.445E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	0.35	1	0.5534

COCHRAN'S Q 0.5349
LARGEST VAR / SMALLEST VAR 1.150
COMPONENT OF VARIANCE FOR BETWEEN GROUPS 6.654
EFFECTIVE CELL SIZE 73.5

VARIABLE	MEAN	SAMPLE SIZE	GROUP STD DEV
CLT	-5.243	74	4.341
ELT	-1.521	73	4.655
TOTAL	-3.395	147	4.500

TABLE 45

ONE WAY ANOVA FOR: CLTM ELTM

SOURCE	DF	SS	MS	F	P
BETWEEN	1	16.73	16.73	0.75	0.3866
WITHIN	145	3.217E+03	22.18		
TOTAL	146	3.234E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	0.08	1	0.7730

COCHRAN'S Q 0.5170
LARGEST VAR / SMALLEST VAR 1.070
COMPONENT OF VARIANCE FOR BETWEEN GROUPS -7.420E-02
EFFECTIVE CELL SIZE 73.5

VARIABLE	MEAN	SAMPLE SIZE	GROUP STD DEV
CLTM	-1.743	74	4.789
ELTM	-1.068	73	4.629
TOTAL	-1.408	147	4.710

TABLE 46

ONE WAY ANOVA FOR: CBT EBT

SOURCE	DF	SS	MS	F	P
BETWEEN	1	23.22	23.22	1.08	0.3002
WITHIN	145	3.115E+03	21.48		
TOTAL	146	3.138E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	2.07	1	0.1506

COCHRAN'S Q 0.5841
LARGEST VAR / SMALLEST VAR 1.404
COMPONENT OF VARIANCE FOR BETWEEN GROUPS 2.366E-02
EFFECTIVE CELL SIZE 73.5

VARIABLE	MEAN	SAMPLE SIZE	GROUP STD DEV
CBT	4.027	74	4.230
EBT	4.822	73	5.012
TOTAL	4.422	147	4.635

TABLE 47

ONE WAY ANOVA FOR: CBTM EBTM

SOURCE	DF	SS	MS	F	P
BETWEEN	1	2.655	2.655	0.14	0.7066
WITHIN	145	2.707E+03	18.67		
TOTAL	146	2.709E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	0.25	1	0.6193

COCHRAN'S Q 0.5292
LARGEST VAR / SMALLEST VAR 1.124
COMPONENT OF VARIANCE FOR BETWEEN GROUPS -2.178E-01
EFFECTIVE CELL SIZE 73.5

VARIABLE	MEAN	SAMPLE SIZE	GROUP STD DEV
CBTM	-3.784E-01	74	4.193
EBTM	-1.096E-01	73	4.446
TOTAL	-2.449E-01	147	4.320

TABLE 48

ONE WAY ANOVA FOR CONTROL DELTA (CDELTA) =

CONTROL SEX (CSEX)

SOURCE	DF	SS	MS	F	P
BETWEEN	1	1.729	1.729	0.04	0.8483
WITHIN	72	3.379E+03	46.92		
TOTAL	73	3.380E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	2.84	1	0.0922

COCHRAN'S Q 0.6463
LARGEST VAR / SMALLEST VAR 1.827
COMPONENT OF VARIANCE FOR BETWEEN GROUPS -1.298
EFFECTIVE CELL SIZE 34.8

CSEX	MEAN	SAMPLE SIZE	GROUP STD DEV
1	3.435	46	7.518
2	3.750	28	5.562
TOTAL	3.554	74	6.850

TABLE 49

ONE WAY ANOVA FOR EXPERIMENTAL DELTA (EDELTA) =

EXPERIMENTAL SEX (ESEX)

SOURCE	DF	SS	MS	F	P
BETWEEN	1	273.9	273.9	4.69	0.0336
WITHIN	71	4.143E+03	58.35		
TOTAL	72	4.417E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	0.56	1	0.4539

COCHRAN'S Q 0.5657
LARGEST VAR / SMALLEST VAR 1.303
COMPONENT OF VARIANCE FOR BETWEEN GROUPS 6.555
EFFECTIVE CELL SIZE 32.9

ESEX	MEAN	SAMPLE SIZE	GROUP STD DEV
1	2.042	48	7.276
2	-2.040	25	8.304
TOTAL	6.438E-01	73	7.639

TABLE 50

ONE WAY ANOVA FOR CONTROL DELTA (CDELTA) =

CONTROL AGE (CAGE)

SOURCE	DF	SS	MS	F	P
BETWEEN	3	83.33	27.78	0.59	0.6277
WITHIN	70	3.297E+03	47.10		
TOTAL	73	3.380E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	2.90	2	0.2346

COCHRAN'S Q 0.7600
LARGEST VAR / SMALLEST VAR 7.161
COMPONENT OF VARIANCE FOR BETWEEN GROUPS -1.736
EFFECTIVE CELL SIZE 11.1

CAGE	MEAN	SAMPLE SIZE	GROUP STD DEV
2	3.529	51	6.827
3	4.250	20	6.077
4	-2.500	2	16.26
5	3.000	1	M
TOTAL	3.554	74	6.863

TABLE 51

ONE WAY ANOVA FOR EXPERIMENTAL DELTA (EDELTA) =

EXPERIMENTAL AGE (EAGE)

SOURCE	DF	SS	MS	F	P
BETWEEN	4	424.8	106.2	1.81	0.1359
WITHIN	68	3.992E+03	58.70		
TOTAL	72	4.417E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	7.25	2	0.0266

COCHRAN'S Q 0.6073
LARGEST VAR / SMALLEST VAR 5.425
COMPONENT OF VARIANCE FOR BETWEEN GROUPS 4.673
EFFECTIVE CELL SIZE 10.2

EAGE	MEAN	SAMPLE SIZE	GROUP STD DEV
2	2.159	44	6.864
3	-3.158	19	10.10
4	2.250	8	4.334
5	-4.000	1	M
6	-2.000	1	M
TOTAL	6.438E-01	73	7.662

TABLE 52

ONE WAY ANOVA FOR CONTROL DELTA (CDELTA) =

CONTROL EDUCATION LEVEL (CEDLVL)

SOURCE	DF	SS	MS	F	P
BETWEEN	6	94.70	15.78	0.32	0.9229
WITHIN	67	3.286E+03	49.04		
TOTAL	73	3.380E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	2.63	2	0.2684

COCHRAN'S Q 0.4758
LARGEST VAR / SMALLEST VAR 2.246
COMPONENT OF VARIANCE FOR BETWEEN GROUPS -4.766
EFFECTIVE CELL SIZE 7.0

CEDLVL	MEAN	SAMPLE SIZE	GROUP STD DEV
1	3.000	1	M
2	9.000	1	M
3	7.000	1	M
4	3.556	45	6.280
5	3.176	17	7.626
6	2.500	8	9.411
7	10.00	1	M
TOTAL	3.554	74	7.003

TABLE 53

ONE WAY ANOVA FOR EXPERIMENTAL DELTA (EDELTA) =

EXPERIMENTAL EDUCATION LEVEL (EEDLVL)

SOURCE	DF	SS	MS	F	P
BETWEEN	5	168.1	33.62	0.53	0.7547
WITHIN	67	4.249E+03	63.41		
TOTAL	72	4.417E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	4.73	4	0.3164

COCHRAN'S Q 0.3299
LARGEST VAR / SMALLEST VAR 18.01
COMPONENT OF VARIANCE FOR BETWEEN GROUPS -3.532
EFFECTIVE CELL SIZE 8.4

EEDLVL	MEAN	SAMPLE SIZE	GROUP STD DEV
2	4.000	2	8.485
3	-5.000E-01	2	2.121
4	0.977	44	8.163
5	2.667E-01	15	9.004
6	4.444E-01	9	4.640
7	-11.00	1	M
TOTAL	6.438E-01	73	7.963

TABLE 54

ONE WAY ANOVA FOR CONTROL DELTA (CDELTA) =

CONTROL EXPERIENCE (CEXP)

SOURCE	DF	SS	MS	F	P
BETWEEN	6	104.9	17.49	0.36	0.9029
WITHIN	67	3.275E+03	48.89		
TOTAL	73	3.380E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	6.74	5	0.2407

COCHRAN'S Q 0.3883
LARGEST VAR / SMALLEST VAR 7.853
COMPONENT OF VARIANCE FOR BETWEEN GROUPS -3.426
EFFECTIVE CELL SIZE 9.2

CEXP	MEAN	SAMPLE SIZE	GROUP STD DEV
1	4.364	11	4.739
2	4.071	28	6.212
4	4.000	1	M
5	4.000	4	10.71
6	1.333	3	13.28
7	5.667	6	6.976
8	2.048	21	7.311
TOTAL	3.554	74	6.992

TABLE 55

ONE WAY ANOVA FOR EXPERIMENTAL DELTA (EDELTA) =

EXPERIMENTAL EXPERIENCE (EEXP)

SOURCE	DF	SS	MS	F	P
BETWEEN	6	1.488E+03	248.0	5.59	0.0001
WITHIN	66	2.929E+03	44.37		
TOTAL	72	4.417E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	8.41	6	0.2096

COCHRAN'S Q 0.2932
LARGEST VAR / SMALLEST VAR 228.7
COMPONENT OF VARIANCE FOR BETWEEN GROUPS 23.46
EFFECTIVE CELL SIZE 8.7

EEXP	MEAN	SAMPLE SIZE	GROUP STD DEV
1	3.500	6	8.735
2	2.394	33	6.485
4	-18.33	3	10.69
5	-1.714	7	7.761
6	-2.500	2	7.071E-01
7	-5.500	4	8.583
8	2.278	18	4.787
TOTAL	6.438E-01	73	6.661

TABLE 56

ONE WAY ANOVA FOR CONTROL DELTA (CDELTA) =

CONTROL EMPLOYMENT (CEMP)

SOURCE	DF	SS	MS	F	P
BETWEEN	6	445.9	74.31	1.70	0.1346
WITHIN	67	2.934E+03	43.80		
TOTAL	73	3.380E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	3.40	4	0.4933
COCHRAN'S Q		0.3332	
LARGEST VAR / SMALLEST VAR		33.81	

COMPONENT OF VARIANCE FOR BETWEEN GROUPS 3.594
EFFECTIVE CELL SIZE 8.5

CEMP	MEAN	SAMPLE SIZE	GROUP STD DEV
1	3.500	36	6.213
2	2.067	15	6.135
3	4.800	10	7.554
4	-14.00	1	M
5	8.000	2	1.414
6	3.000	1	M
8	5.889	9	8.223
TOTAL	3.554	74	6.618

TABLE 57

ONE WAY ANOVA FOR EXPERIMENTAL DELTA (EDELTA) =

EXPERIMENTAL EMPLOYMENT (EEMP)

SOURCE	DF	SS	MS	F	P
BETWEEN	6	339.7	56.61	0.92	0.4898
WITHIN	66	4.077E+03	61.77		
TOTAL	72	4.417E+03			

AT LEAST ONE GROUP VARIANCE IS NEAR ZERO;
VARIANCE-EQUALITY TESTS CANNOT BE COMPUTED.
COMPONENT OF VARIANCE FOR BETWEEN GROUPS -5.862E-01
EFFECTIVE CELL SIZE 8.8

EEMP	MEAN	SAMPLE SIZE	GROUP STD DEV
1	1.848	33	6.915
2	2.857E-01	14	10.46
3	-4.200	10	9.151
4	1.143	7	4.598
5	-2.000	2	0.000
6	5.000	1	M
8	2.500	6	6.979
TOTAL	6.438E-01	73	7.860

TABLE 58

ONE WAY ANOVA FOR CONTROL DELTA (CDELTA) =

CONTROL GRAPH USE (CUSE)

SOURCE	DF	SS	MS	F	P
BETWEEN	5	509.2	101.8	2.41	0.0447
WITHIN	68	2.871E+03	42.22		
TOTAL	73	3.380E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	10.77	4	0.0292

COCHRAN'S Q 0.3584
LARGEST VAR / SMALLEST VAR 10.43
COMPONENT OF VARIANCE FOR BETWEEN GROUPS 5.530
EFFECTIVE CELL SIZE 10.8

CUSE	MEAN	SAMPLE SIZE	GROUP STD DEV
1	7.000	1	M
2	12.00	4	9.416
3	4.412	17	8.024
4	1.258	31	5.657
5	5.364	11	7.339
6	3.500	10	2.915
TOTAL	3.554	74	6.498

TABLE 59

ONE WAY ANOVA FOR EXPERIMENTAL DELTA (EDELTA) =

EXPERIMENTAL GRAPH USE (EUSE)

SOURCE	DF	SS	MS	F	P
BETWEEN	5	685.7	137.1	2.46	0.0412
WITHIN	67	3.731E+03	55.69		
TOTAL	72	4.417E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	22.26	5	0.0005

COCHRAN'S Q 0.5483
LARGEST VAR / SMALLEST VAR 27.47
COMPONENT OF VARIANCE FOR BETWEEN GROUPS 7.202
EFFECTIVE CELL SIZE 11.3

EUSE	MEAN	SAMPLE SIZE	GROUP STD DEV
1	5.444	9	4.927
2	-3.714	7	15.30
3	9.333E-01	15	5.837
4	-1.852	27	6.747
5	4.833	6	8.976
6	3.444	9	2.920
TOTAL	6.438E-01	73	7.462

TABLE 60

ONE WAY ANOVA FOR CONTROL DELTA (CDELTA) = CONTROL GRAPH CONSTRUCTION (CCONST)

SOURCE	DF	SS	MS	F	P
BETWEEN	5	97.77	19.55	0.41	0.8447
WITHIN	68	3.283E+03	48.27		
TOTAL	73	3.380E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	4.78	4	0.3110

COCHRAN'S Q 0.3721
LARGEST VAR / SMALLEST VAR 4.659
COMPONENT OF VARIANCE FOR BETWEEN GROUPS -3.311
EFFECTIVE CELL SIZE 8.7

CCONST	MEAN	SAMPLE SIZE	GROUP STD DEV
1	7.000	1	M
2	6.500	2	3.536
3	5.364	11	6.682
4	2.822	45	7.632
5	3.222	9	4.295
6	4.667	6	4.761
TOTAL	3.554	74	6.948

TABLE 61

ONE WAY ANOVA FOR EXPERIMENTAL DELTA (EDELTA) = EXPERIMENTAL GRAPH CONSTRUCTION (ECONST)

SOURCE	DF	SS	MS	F	P
BETWEEN	5	77.11	15.42	0.24	0.9429
WITHIN	67	4.340E+03	64.77		
TOTAL	72	4.417E+03			

	CHI SQ	DF	P
BARTLETT'S TEST OF EQUAL VARIANCES	6.61	4	0.1580

COCHRAN'S Q 0.3721
LARGEST VAR / SMALLEST VAR 3.648
COMPONENT OF VARIANCE FOR BETWEEN GROUPS -5.486
EFFECTIVE CELL SIZE 9.0

ECONST	MEAN	SAMPLE SIZE	GROUP STD DEV
1	3.000	1	M
2	-3.000	2	7.071
3	1.176E-01	17	11.03
4	5.366E-01	41	6.686
5	1.167	6	8.796
6	3.167	6	5.776
TOTAL	6.438E-01	73	8.048

Bibliography

American National Standards Institute. *Time Series Charts* (ANSI Y15.2M), New York: The American Society of Mechanical Engineers, 1979.

Analytical Software. *Statistix User's Manual (Version 3.5)* St. Paul: June 1991.

Auger, B. Y. *How to Run Better Business Meetings,* Minnesota Mining and Manufacturing Company, 1979.

Babbitt, Bettina A. and Charles O. Nystrom. *Questionnaire Construction Manual Annex, Questionnaires: Literature Survey and Bibliography.* Research Product 89-21. U.S. Army Research Institute for the Behavioral and Social Sciences, Field Unit at Fort Hood TX. June 1989.

Christensen, Maj David S. and Capt Albert Larkin. "Criteria for High Integrity Graphs," *Journal of Managerial Issues,* Volume IV, Number 1: 130-153 (Spring 1992).

Cleveland, William S. *The Elements of Data Graphing,* Monterey CA: Wadsworth Advanced Books and Software, 1985.

Devore, Jay L. *Probability and Statistics for Engineering and the Sciences.* Pacific Grove CA: Brooks/Cole Publishing Co., 1991.

Emory, William C. and Donald R. Cooper. *Business Research Methods (Fourth Edition).* Boston: Irwin, 1991.

Howard, Bill and Gerald Kunkel. "Designing Great Graphics," *PC Magazine,* Volume 7, Number 16: 92-95+ (September 27, 1988).

Johnson, Johnny R. and others. "Pictures that Lie: The Abuse of Graphics in Annual Reports," *Management Accounting 10:* 50-56 (October 1980).

Joint Committee on Standards for Graphic Representation. Preliminary Report Published for the Purpose of Inviting Suggestions for the Benefit of the Committee." *Journal of the American Statistical Association, 112*: 90-97 (December 1915).

Kern, Capt Charles W. *Can Graphs Mislead Decision Makers When Formulated in Violation of Tufte's Lie Factor.* MS thesis, AFIT/GCA/LSY/91S-5. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1991 (AD-A243944).

Larkin, Capt Albert A. *Misleading Graphics: Can Decision Makers Be Affected By Their Use?* MS thesis, AFIT/GSM/LSY/90S-18. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1990 (AD-A229621).

Lefferts, Robert. *Elements of Graphics: How to Prepare Charts and Graphs for Effective Reports,* New York: Harper and Row, 1981.

MacGregor, A.J. *Graphics Simplified,* Toronto: University of Toronto Press, 1979.

Nunnally, Jum C. *Psychometric Theory (Second Edition)*. New York: McGraw-Hill Book Co., 1978.

Reynolds, Daniel E. Class notes from MATH 592, Managerial Statistics I. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, July 1991.

Rogers, Anna C. *Graphics Charts Handbook,* Washington: Public Affairs Press, 1961.

Schmid, Calvin F. and Stanton E. Schmid. *Handbook of Graphic Representation,* New York: John Wiley and Sons, 1954.

Schmid, Calvin F. *Statistical Graphics Design Principles and Practices,* New York: John Wiley and Sons, 1983.

Seymour, Jim. "At the Top of the Charts," *PC Magazine, Volume 7, Number 5:* 93-96+ (March 15, 1988).

Shane, Guy S. Associate Professor of Organizational Behavior and Management. Personal Interview. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, May 1992.

Spear, Mary Eleanor. *Practical Charting Techniques,* New York: McGraw Hill Book Company, 1969.

Steinbart, Paul John. "The Auditor's Responsibility for the Accuracy of Graphs in Annual Reports: Some Evidence of the Need for Additional Guidance," *Accounting Horizons 18:* 60-70 (October 1986).

Tan, Joseph K. H. and Izak Benbasat. "Processing of Graphical Information: A Decomposition Taxonomy to Match Data Extraction Tasks and Graphical Representation," *Information Systems Research, Volume 1, Number 4:* 416-439 (December 1990).

Taylor, Barbara G. *The Potential Impact of Graphic Formatting on the Reliability of Financial Reporting,* Dissertation, Texas Tech University, December 1983.

Tufte, Edward R. *The Visual Display of Quantitative Information,* Cheshire CT: Graphics Press, 1983.

Vita

Captain Clark R. Carvalho was born in Ann Arbor, Michigan, and graduated from high school in Salt Lake City, Utah in 1976. He enlisted in the Air Force in 1978, where he was trained as a Nuclear Weapons Specialist, with subsequent assignments at Minot AFB, North Dakota and Ramstein AB, Germany. In 1983, he was selected for the Airman's Education and Commissioning Program, where he received his bachelors degree in Electrical Engineering from Northern Arizona University in 1985. After receiving his commission in April, 1986, he was assigned to the Rome Air Development Center, Hanscom AFB, MA as an Electrical Engineer. In 1990, he received a masters degree in Engineering Management from Western New England College. In April, 1989 he was transferred to the Electronic Systems Division, Airspace Management Program Office where he was program manager for the Digital Bright Radar Indicator Tower Equipment program and Chief of the Air Traffic Control and Landing Systems Division until entering the School of Systems and Logistics, Air Force Institute of Technology, in May, 1991.

Permanent Address: 5310 Access Rd
Dayton, OH 45431

Vita

Captain Michael D. McMillan was born May 29, 1956 in Pine Hurst, North Carolina. He graduated from Terry Sanford High School in Fayetteville, North Carolina in 1974. After high school he attended North Carolina State University and Appalachian State University, where he graduated with a bachelors degree in Business Administration in 1980. He enlisted in the United States Air Force and received his commission through Officer Training School in 1981. Upon commissioning as a second lieutenant, he entered Undergraduate Navigator Training and later was assigned to Bergstrom AFB, Texas for upgrade training in the RF-4C. At the completion of flying training he was assigned to the 26th Tactical Reconnaissance Wing (TRW), Zweibrucken AB, Germany in 1983 for a two year operational tour. This assignment was followed by selection as an instructor weapon systems officer for RF-4C formal upgrade training at Bergstrom AFB. During this tour he served as squadron scheduler, life support officer, and assistant chief of the RF-4C Operations Training Division (OTD). In his OTD capacity he was involved in the total revision of all RF-4C formal training programs and the design and procurement of an upgraded RF-4C flight simulator. In 1989 he was again assigned to the 26th TRW and held positions as instructor, chief of current operations, flight commander, and temporary squadron operations officer. In May 1991 he entered the Systems Management masters degree program at the School of Systems and Logistics, Air Force Institute of Technology.

Permanent Address: 3405 Rogers Dr.
Fayetteville, NC 28303

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (*Leave blank*)

2. REPORT DATE
September 1992

3. REPORT TYPE AND DATES COVERED
Master's Thesis

4. TITLE AND SUBTITLE
GRAPHIC REPRESENTATION IN MANAGERIAL DECISION MAKING: THE EFFECT OF SCALE BREAK ON THE DEPENDENT AXIS

5. FUNDING NUMBERS

6. AUTHOR(S)
Clark R. Carvalho, Capt, USAF
Michael D. McMillan, Capt, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology, WPAFB OH 45433-6583

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GSM/LSY/92S-5

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)
This thesis investigated whether a scale break on the dependent axis of a graph affects a decision maker's interpretation of the data presented in the graph. A literature search revealed criteria for constructing high integrity graphs and formatting scale breaks. An experiment was conducted to determine the effect of a scale break on the dependent axis. Graphs following the criteria for high integrity graphs were presented to the control group, while graphs following the criteria for high integrity graphs, with the exception of the scale break, were presented to the experimental group. Using a parametric two-sample t test and a non-parametric Rank Sum test, it was shown that data presented in a graph with a scale break was interpreted differently from data presented in a graph without a scale break. Analysis of variance was conducted on the demographic factors for each subject. In the experimental group, sex and professional experience were factors that led to different interpretations of graphs with a scale break. The level of experience using graphs in decision making was also a factor that led to different interpretations of graphs in both groups.

14. SUBJECT TERMS
Decision Making, Misleading, Graphics, Graph,

15. NUMBER OF PAGES
133

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

AFIT Control Number AFIT/GSM/LSY/92S-5

AFIT RESEARCH ASSESSMENT

The purpose of this questionnaire is to determine the potential for current and future applications of AFIT thesis research. Please return completed questionnaires to: AFIT/LSC, Wright-Patterson AFB OH 45433-9905.

1. Did this research contribute to a current research project?

   a. Yes b. No

2. Do you believe this research topic is significant enough that it would have been researched (or contracted) by your organization or another agency if AFIT had not researched it?

   a. Yes b. No

3. The benefits of AFIT research can often be expressed by the equivalent value that your agency received by virtue of AFIT performing the research. Please estimate what this research would have cost in terms of manpower and/or dollars if it had been accomplished under contract or if it had been done *in-house*.

   Man Years ______________ $ ______________

4. Often it is not possible to attach equivalent dollar values to research, although the results of the research may, in fact, be important. Whether or not you were able to establish an equivalent value for this research (3, above) what is your estimate of its significance?

   a. Highly Significant b. Significant c. Slightly Significant d. Of No Significance

5. Comments

______________________ ______________________

Name and Grade Organization

______________________ ______________________

Position or Title Address

(Fold down on outside — seal with tape)

DEPARTMENT OF THE AIR FORCE
AFIT/ LSC
WRIGHT-PATTERSON AFB OH 45433-6583

OFFICIAL BUSINESS

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL

FIRST CLASS MAIL PERMIT NO 1006 DAYTON OH

POSTAGE WILL BE PAID BY THE ADDRESSEE

WRIGHT-PATTERSON AIR FORCE BASE

AFIT/ LSC
WRIGHT-PATTERSON AIR FORCE BASE
DAYTON OH 45433-9905

FOLD IN